REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Ferhat Abbas Sétif-1

Faculté des Sciences

Département d'Informatique

THÈSE

Thèse de doctorat en cotutelle de l'Université Ferhat Abbas Setif-1 et de l'Université Bourgogne Franche-Comté

Présentée par

Mohamed Sofiane BATTA

Pour l'obtention du diplôme de

Doctorat 3ème Cycle LMD en Informatique

Option : Sureté de fonctionnement dans les réseaux dynamiques

Thème

Gestion Efficace des Canaux de Communication dans l'Internet des Objets

Soutenue le: 04 Janvier 2023 Devant le jury composé de:

Pr. Francois SPIES	Professeur, Université Bourgogne Franche-Comté	Président
Pr. Zibouda ALIOUAT	Professeur, Université de Ferhat Abbas Sétif-1	Directrice de thèse
Pr. Nabila LABRAOUI	Professeur, Université de Tlemcen	Rapporteur
Dr. Hakim MABED	MCF, Université Bourgogne Franche-Comté	Directeur de thèse
Dr. Benoit HILT	MCF, Université de Haute-Alsace	Rapporteur
Dr. Chirihane GHERBI	MCA, Université de Ferhat Abbas Sétif-1	Examinateur

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

Ministry of Higher Education and Scientific Research

University of Ferhat Abbas Setif-1

Faculty of Sciences

Department of Computer Science

Université Ferhat Abbas Sétif 1

THESIS

Doctoral thesis jointly supervised by Ferhat Abbas Setif-1 University and Bourgogne Franche-Comté University

Presented by

Mohamed Sofiane BATTA

As a requirement to aim for the degree of

Doctorate 3rd Cycle LMD in Computer Science

Option: Dependability in dynamic networks

Title

Efficient Management of Communication Channels in IoT Networks

Defended on: January 04, 2023 Board of examiners:

Pr. Francois SPIES	Professor, University of Bourgogne Franche-Comté	Chairman
Pr. Zibouda ALIOUAT	Professor, University of Ferhat Abbas Sétif-1	Supervisor
Pr. Nabila LABRAOUI	Professor, University of Tlemcen	Rapporteur
Dr. Hakim MABED	MCF, University of Bourgogne Franche-Comté	Supervisor
Dr. Benoit HILT	MCF, University of Haute-Alsace	Rapporteur
Dr. Chirihane GHERBI	MCA, University of Ferhat Abbas Sétif-1	Examiner

Acknowledgment

First of all, I thank God for the strength and blessing he gives me to accomplish this work.

I would like to express my heartfelt thanks and sincere gratitude to my thesis supervisors **Pr. Zibouda ALIOUAT** and **Dr. Hakim MABED** for their insightful instructions during the development of this thesis. They provided me with a boundless amount of support, motivation, openness, guidance, fruitful discussions and valuable advices throughout the course of my PhD journey.

I would like to express my gratitude for all the jury who honored me by accepting to be part of thesis jury. I would like to thank Pr. Nabila LABRAOUI and Dr. Benoit HILT who devoted their valuable time to review this manuscript. I would also like to thank Pr. Francois SPIES and Dr. Chirihane GHERBI for their precious work as examiners.

I am very grateful and I would like to express my deep gratitude to Pr. Saad HAROUS for his impressive availability, generosity and valuable remarks that helped me in better improving the work done during the entire thesis.

I express my sincere gratitude to all my family for the valuable support they provide to me. Especially, I am deeply grateful to my mother, my father and my brothers for their encouragement and continued support, which has helped me stay strong and focused on the thesis work.

I extend my sincere thanks to all the members of the Department of Computer sciences (UFAS) and the DISC FEMTO-ST (UBFC) which gave me an excellent work environment as well as an extremely pleasant stay.

Last but not least, I would like to thanks all my friends for their affection and encouragement and for each person who contributed directly or indirectly in the successful outcome of this work.

List of publications

This thesis gave rise to several publications. Here is the list:

Journal articles

1. M. S. Batta, H. Mabed, Z. Aliouat, S. Harous, “Battery State-of-Health Prediction based Clustering for Lifetime Optimization in IoT Networks”. IEEE Internet of Things Journal, 2022. Impact Factor: 10.23 (Q1) [1].

2. M. S. Batta, H. Mabed, Z. Aliouat, S. Harous, “A distributed multi-hop intra-clustering approach based on neighbors two-hop connectivity for IoT networks”. Sensors, 21(3), 873, 2021. Impact Factor: 3.84 (Q1) [2].

3. M. MERAH, Z. Aliouat, Y. Harbi, M. S. Batta, “Machine Learning-based Clustering Protocols for IoT Networks: Overview and Multi-directional Challenges”, Computers & Electrical Engineering (Major revision).

4. M. S. Batta, Z. Aliouat, H. Mabed, M. MERAH, “A Dynamic Unequal Clustering algorithm based Lifetime Optimization for Urban IoT Networks”, International Journal of Communication Systems (Under review).

Conference articles

1. M. S. Batta, Z. Aliouat, H. Mabed, M. MERAH, “An Improved Lifetime Optimization Clustering using Kruskal’s MST and Batteries Aging for IoT Networks", The International Symposium on Networks, Computers and Communications (ISNCC), Shenzhen, China, July 2022 [3].

2. M. S. Batta, H. Mabed, Z. Aliouat, “State of Health Optimization Based Unequal Clustering in IoT Networks”, In International Wireless Communications and Mobile Computing Conference (IWCMC), Dubrovnik, Croatia, June 2022 [4].

3. M. S. Batta, Z. Aliouat, H. Mabed, M. MERAH, “A Distributed Energy-Efficient Unequal Clustering based Kruskal Heuristic for IoT Networks”, In 14th International

Conference on Applications of Graph Theory in Wireless Ad hoc Networks and Sensor Networks (GRAPH-HOC), Copenhagen, Danemark, 2022 [5].

4. M. MERAH, Z. Aliouat, M. S. Batta, "A Hybrid Neural Network and Graph theory based Clustering Protocol for dynamic IoT Network", International Conference on Advanced Aspects of Software Engineering (ICAASE), Constantine, Algerie, 2022 [6].

5. H. Mabed, M. S. Batta, Z. Aliouat, “Optimization of rechargeable batteries lifespan in wireless networking protocols”, In Mobile and Ubiquitous Systems (MobiQuitous), Madrid, Spain, 2020 [7].

6. M. S. Batta, Z. Aliouat, H. Mabed, S. Harous, "LTEOC: Long Term Energy Optimization Clustering For Dynamic IoT Networks", In IEEE Ubiquitous Computing, Electronics & Mobile Communication Conference, USA, 2020 [8].

7. M. S. Batta, H. Mabed, Z. Aliouat, “Dynamic clustering based energy optimization for IoT network”, In International Symposium on Modelling and Implementation of Complex Systems, pp. 77–91, Algeria, 2020 [9].

8. M. S. Batta, Z. Aliouat, S. Harous, “A distributed weight-based TDMA scheduling algorithm for latency improvement in IoT” in IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), pp. 768–774, USA, 2019 [10].

9. M. S. Batta, S. Harous, L. Louail, and Z. Aliouat, “A distributed TDMA scheduling algorithm for latency minimization in internet of things”, In IEEE International Conference on Electro Information Technology (EIT), pp. 108–113, USA, 2019 [11].

Contents

List of Figures

LIST OF TABLES

General Introduction

With the massive growth of smart devices and the rapid progress of wireless communication technologies, the internet of things (IoT) witnesses a significant evolution and received the attention of both industrial and academic fields during the last decades. Undoubtedly, the main strength of the IoT technology is its high impact on several aspects of users' daily-life. From a private user perspective, the IoT paradigm will be visible in both working and domestic fields (assisted living, e-healthcare, enhanced learning, etc). From a business user point of view, the IoT effect rises in several industrial fields, such as automation and manufacturing, logistics, intelligent transportation systems, business management, and processing [17]. Furthermore, this evolution invaded the human's daily life services in both public (smart cities, smart buildings, smart healthcare, agriculture, smart lighting, etc.) and private areas (smart homes, smart factories). The IoT refers to a paradigm where a myriad number of smart and tiny devices (e.g. smart phones, sensors, actuators, home appliances, etc) are connected to the internet and communicate in an intelligent fashion to achieve a common goal.

The miniaturization of the electronic devices arises from the fact that the recent electronic components (processors, storage, radio antennas and battery capacities) become available with a low cost and size. Moreover, the integration of these devices to modern networks has been made possible by the development of different networking protocols (e.g. RPL , CoAP, 6LoWPAN, etc.) and several communication technologies (e.g. BLE, Zigbee, NFC, etc.) [18]. This trend promotes the utilization of these devices into different systems such as the wireless sensors networks (WSN) [19], Radio-Frequency IDentification (RFID), etc. [17].

IoT provides many specific advantages and characteristics, such as the integration of several internet services. For example, Cloud Computing (CC) technology can be integrated to allow the application's data to be stored in the cloud, so that they can be monitored by the user terminal with low capacity [15]. Moreover, these modern networks have the ability to integrate multiple objects and combine multiple networks using different communication technologies [20; 21]. This latter allows the exploitation of different wireless communication technologies, which facilitates the network scalability with a significant number of connected devices.

0.0.1/ Context and Motivations

Even though IoT has many advantages, it faces a number of challenges [22]. The small scale aspect of the electronic devices exposes the network to many issues. These issues are mainly related to the restricted physical capabilities (reduced processing power, low transmission range and small energy). Firstly, among the main challenges, the integration of the rapid growing number of devices into the IoT ecosystem. Secondly, the exploitation of the adequate communication protocols that ensure the connectivity under resource restricted system. Thirdly, the investigation and the development of new efficient protocols that consider the implementation problems that face these modern system (e.g. energy efficiency, security, reliability, latency, communication channel management, routing, network congestion and collisions, etc) in order to ensure the efficiency of the system and prolong the network lifetime [23]. Finally, IoT networks face several challenges that already exist in conventional networks, such as the dynamic topology, variable network density, communication models and network fragmentation. The creation of such a smart network requires, in one form or another, the connection of all the IoT devices to the internet. This connection generates a massive amounts of data that need to be relayed and an enormous number of wireless connections that have to be established [15].

Research and development activities on wireless communication technology constitute a critical segment of research and development in telecommunication area. Communication resources such as channel bandwidth, radio spectrum and transmission characteristics represent a fundamental component for any wireless network. Therefore, communication resource management is an important topic for the development of modern wireless networks. Efficient allocation and management of communication channels in a dynamic environment with different devices requirement is essential for the practical deployment of wireless communication systems [24].

Considering all the above issues, network management becomes an attractive topic in modern wireless networks to reach the best performance and achieve a high-level of availability. Nevertheless, the realization of such a performance objective is not straightforward, due to the intrinsic characteristics of IoT's restricted resources network. Moreover, the choice of low power mechanism for collecting and efficiently routing the data in the networks is still relevant, especially with the expendability aspect of such networks). Therefore, the design and maintenance of this dynamic system requires an efficient management strategy and a scalable architecture. In this context, literature researches have turned to the use of new network structuring techniques for the management of the large network and addressing the system challenges to provide appropriate conditions to run the different applications.

Network clustering is one of the widely used techniques in conventional wireless networks for efficiently managing the network. Its considerably strengthen the network performances compared to the classic structure by considering numerous applications such as topology management, overhead minimization, routing task, data aggregation, scalability, quality of services (QoS), channel access control. Therefore, clustering is a keystone technique for the design and modeling of IoT network protocol and applications in order to ensure the load balancing, network lifetime extinction, stable structure and channels bandwidth exploitation. However, some performance attribute are not well considered by the existing network management solutions such as the scalability, self-stabilization and the long-term energy preservation [25].

Clustering techniques are diverse [26; 27; 23; 28; 29; 30; 2; 31; 32; 33], the common approach consist of grouping the nodes of the network into clusters with a representative node in each cluster responsible of collecting and aggregating the collected information. Communications within a wireless network are provided through a layered model used by all network elements. In this model (typically inspired from the Open Systems Interconnection model [34]), the layers are implemented independently of each other, each layer being responsible for ensuring a specific features to optimize certain metrics of the system.

0.0.2/ Contributions

In this thesis, we address the problem of communication and topology management in IoT networks and propose efficient solutions to meet the requirements of such networks and improve their durability. Therefore, the general objective of this thesis focuses on the design, modeling and simulation of new management protocols for the large scale network. More precisely, we focused on developing protocols for the MAC and routing layer since these two layers are responsible for controlling the communication access and managing the network topology. We propose different management approaches for enhancing the performances of the network and confront their challenges.

As the minimization of communication latency is one of the crucial constraints requested in wireless networks [35]. A sharing knowledge between the routing layer and the MAC layer protocols is required. We, firstly, address the cross layer design protocols to overtake the latency constraint (chapter 3). We propose a fully distributed TDMA scheduling approache for the MAC layer that uses the routing protocol information's. As a contribution, we introduced a distributed weight based TDMA scheduling called WB-DTSA [10]. The idea is to allocate transmission slots to devices according to their position in the routing tree and their distance toward the base station. The objective of the proposal is to improve the TDMA scheduling efficiency, reduce the transmission latency and the energy wastage.

Secondly, we focus on the topology management problem, more precisely, we apply the clustering technique to improve the performance of a large scale networks (chapter 4). We perceived a restricted power devices aspect and we introduced new multi-hop clustering techniques for large and dense networks. The contribution proposes a distributed approach for the k-hop intra-clustering called Distributed Clustering based 2-Hop Connectivity (DC2HC) [2]. The algorithm uses the two-hop neighbor's connectivity to elect the appropriate set of cluster heads and strengthen the clusters connectivity. Its considers a self stabilisation technique to tolerate the network dynamic topology. The objective is to optimize the set of cluster heads to minimize the number of long range transmissions which reduce the energy harvesting and expand the network lifetime.

Thirdly, we considered a rechargeable aspect of devices batteries and we focus on the long-term energy optimization of the network lifespan by considering the state of health (SOH) of the devices' batteries instead of using the conventional State of Charge (SoC) (chapter 5). In this concern, we present a new dynamic clustering technique using an original long-term vision of energy optimization for IoT network (LTEOC: Long-Term Energy Optimization Clustering) [8], by taking into consideration the rechargeable battery aging aspect during the clustering process to increase the network durability in the long-term.

Afterward, in chapter 6, we proposed a novel SOH prediction based Clustering Approach for Long-term Energy optimization ($LECA\_SOH$) [1]. In this latter, instead of directly using the rechargeable battery SOH in the clustering process. We rather use an original metric which predict the forthcoming SOH of the batteries before the election of the CHs. In this case, nodes that risk less potential battery degradation during the next round are selected, which in turn extends the system life significantly.

0.0.3/ Dissertation outline

This thesis is divided into two main parts: background and contributions. The dissertation is organized as follows.

In the background part:

- **Part One: Background**
  - **Chapter 1**: give background knowledge about the IoT networks. It provides an overview on IoT networks and its component. Followed by a presentation of the architecture, characteristics and communication technologies. Then, it exhibits some applications and provides a presentation of the requirements and the challenges facing such networks.
  - **Chapter 2**: present the state of the art on topology and communication management solution for wireless networks in both MAC and routing sub layers. It

is divided into two main parts: The first part presents the state-of-the art on MAC and cross layer protocols. The second part provides a state-of-the-art on routing layer protocols. In particular, we provide a classification of existing clustering algorithms for wireless networks.

- **Part Two: Contributions** this chapter is devoted to the presentation of the different contributions, it is devided 7 chapters spread into three sections.
  - **Section 1**: this section address the cross layer design protocols to overtake the latency constraint
    - **Chapter 3**: we present a new distributed weight-based TDMA-MAC scheduling approach (WB-DTSA) for IoT networks [10]. The proposed solution takes advantage of routing information to improve the communication latency and the schedule length.
  - **Section 2**: this section, we present the details of research contributions concerning the topology clustering, which are organized as follows.
    - **Chapter 4**: we proposes a distributed approach for the k-hop intra-clustering called Distributed Clustering based 2-Hop Connectivity (DC2HC) [2]. The approach uses the multi-hop topology knowledge to elect the appropriate set of cluster heads and improve the clusters connectivity.
  - **Section 3**: this section illustrates the works concerning the the long-term energy optimization of the network lifespan, which are organized as follows.
    - **Chapter 5**: we exhibits a new dynamic clustering technique using an original long-term vision of energy optimization for IoT network (LTEOC) [8]. The approach takes into consideration the rechargeable battery aging aspect to extend the network lifetime in the long-term.
    - **Chapter 6**: proposes a novel SOH prediction based Clustering Approach for Long-term Energy optimization ($LECA\_SOH$) [1]. This latter uses an original metric which predict the forthcoming SOH of devices battery before the clusters formation, which significantly extends the system life.
- **Finally**, we conclude our thesis with a general discussion and future perspectives of the work.

I

State of the Art

1

BACKGROUNG

1.1/ INTRODUCTION

The generalization of mobile telecommunication devices (smartphones, laptops and tablets) and the growing use of Internet services (web, e-mail, video on demand, and social networks) along with the availability of the appropriate communication protocols standards have led to the emergence of the Internet of Things (IoT) paradigm.

IoT domain has experienced a continuous interest from scientific and industrial community due to their applications in our daily life (smart homes, health monitoring, smart cities, smart agriculture, etc.). IoT devices are integrated into many kind of objects, such as human bodies, clothes, houses, campuses, buildings and production engines. This use permits the emergence of interesting services that cover diverse domains and allows the creation of a modern smart-objects' environment. These smart systems are able to predict and understand human behaviors and how to react to any kind of stimuli. The application area of IoT networks are numerous and still growing. They cover environmental monitoring, living assistance, target tracking, e-health, critical infrastructures monitoring and homeland security, etc. The low-cost, easy deployment and vast usage of WSNs have contributed immensely to these growths [15].

An IoT network can be seen as a multitude of autonomous telecommunication devices that collaborate together to accomplish a specific task including the interact with human users [17]. According to [12], the number of Internet connected objects had surpassed the earth's human population [15], the global number of connected devices grew from 3.6 billion in 2015 to 13.8 billion in 2021, and it is expected to grow to 30.9 billion in 2025 (figure 1.1). It is legitimate to consider IoT networks as one of the facets of the next Internet generation [36], also known as the Internet of Everything. In this chapter, we provide a background knowledge about the IoT networks technology.

In spite of the growth potential of the IoT network and the high emergence of it attractive services, it faces many challenges, mainly related to the communications, routing,

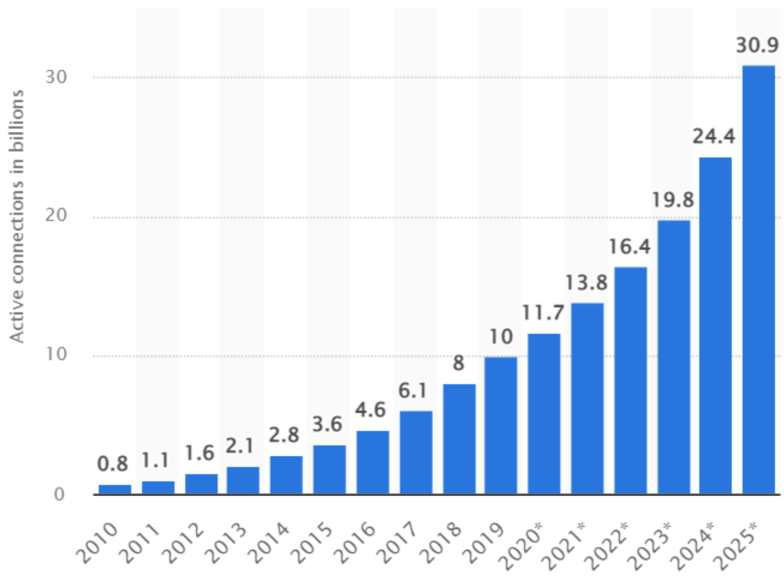


Figure 1.1: Growth of connected IoT devices 2010-2025 [12]

protocols standardization and the restricted devices capabilities (energy, processing and storage). As IoT networks are deployed in large scale environments, managing a large number of mobile heterogeneous devices is a real issue. Therefore, the availability, reliability and interoperability constitute a major concern that affects the network efficiency. The rest of the chapter introduces the fundamental concepts of IoT and their architecture, component, applications, communication technologies, and challenges.

1.2/ IoT DEFINITION: *One paradigm, many visions*

The basic idea of the IoT concept is the connection of pervasive things, or objects, present around us (such as Radio-Frequency IDentification (RFID) tags, sensors, laptops, actuators, smartphones, etc.) with the Internet. These connected heterogeneous objects interact and cooperate together to provides several attractive services that promote the users daily life [18].

The first definition of IoT arises from a "*Things oriented*" context where the considered "*Things*" were represented by simple items: Radio-Frequency IDentification (RFID [1]) tags. Indeed, the terms "*Internet of Things*" was, initially, devoted to the Auto-ID Labs [2]. The

[1] RFID: a technology that promotes the automatic identification and tracking of tags attached to any object or people chips [17]. They are made of a microchip (also called a label or tag) and an antenna which communicate by radio waves with a reader over distances that can range from a few centimeters to several tens of meters.

[2] Auto-ID Labs: a worldwide network of academic research laboratories in the area of networked RFID and the emerging sensing technologies (https://www.autoidlabs.org/).

main objective of the Auto-ID Labs together with EPCglobal [3] is to architect the IoT vision. To reach this purpose, these institutions make their primary focus on the development of the Electronic Product Code (EPC) to promote the expansion use of RFID in worldwide modern trading networks in order to produce an industry-driven global standards for the EPCglobal Network. These standards are mainly designed to improve the traceability of objects. This is undoubtedly not the only technique to support the deployment of the IoT paradigm. However, it constitutes a key component that encourages the development of the IoT vision.

The International Telecommunication Union (ITU) [4] describe the Internet of things as *the global infrastructure that enables advanced services by interconnecting physical and virtual things using evolving interoperable information and communication technologies.*

The European Research Cluster on the Internet of Things - IERC [5] defines the IoT as *a dynamic universal network infrastructure dotted with a self-configuration ability built on standard and scalable communication protocols. The physical and virtual 'things' that constitute the network have identities, physical attributes, virtual personalities, intelligent interfaces, and are fully integrated into the information network. These 'Things' represent an active participant in business, information and social processes. They are enabled to communicate among themselves and interact with the environment by exchanging the collected data. They react autonomously to potential events and are able to trigger a behavior and create services with or without direct human's intervention*.

Without doubt, the main power of the IoT idea is the significant impact it will have on several sides of our quotidian life and the behavior of potential users. It can be concluded that there are different definitions of the Iot vision, which diverge according to the point of view of researchers and their own investigation areas. However, All those definition converge to a common viewpoint that gives a concise definition of the IoT paradigm. In Figure 1.2, the primary concepts, technologies and standards are classified with reference to the IoT vision. It obviously appears that the IoT paradigm is the result of the convergence of all the main visions addressed above.

1.3/ IoT devices

The wireless IoT devices represent small electronic devices equipped with sensing, computation, actuating and communication capabilities, they are characterized by a limited

[3] EPCglobal: an innovative organization established to achieve worldwide adoption and standardization of Electronic Product Code (EPC) technology (https://www.gs1.org/epcglobal).

[4] ITU: a specialized agency of the United Nations responsible for all matters related to information and communication technologies (https://www.itu.int/).

[5] IERC: established as part of Europe's ambition to shape a future Internet of Things for its businesses and citizens (http://www.internet-of-things-research.eu).

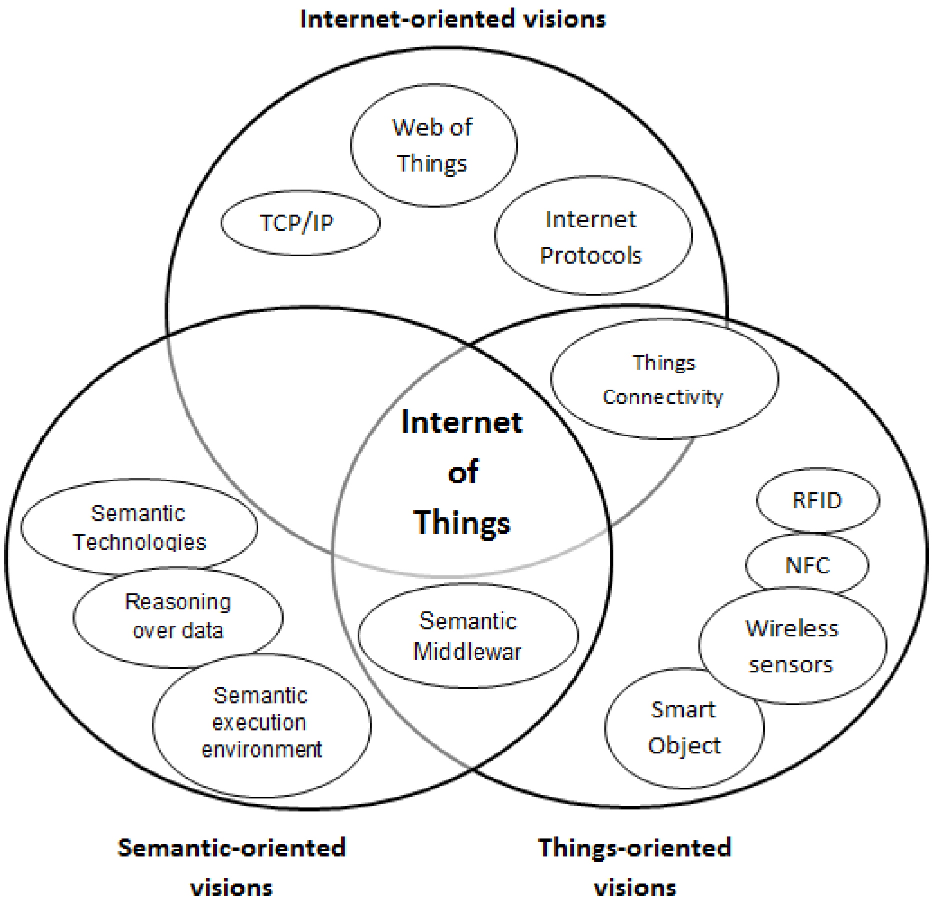


Figure 1.2: The Internet of Things paradigm with different visions

power supply along with some additional application dependent functionalities such as localization system. Sometimes, the IoT device corresponds to a sensor that measures physical phenomena of the surrounding environment and forwards the gathered data to a distant controller. We talk then about Wireless Sensor Network (WSN). Additionally, these devices may be equipped with actuators that interpret the control signals generated from the sensing unit and perform appropriate actions to cope with their prompt environment alteration. The analysis of the sensed data may be carried with a partial human assistance or completely handled by a remote service (e.g. Cloud Computing services, fog networking, etc.). The low-power incorporated circuits and the wireless communications promote the development of these miniature devices and their usage in remote sensing applications as they fill the gap between the physical and the electronic world [37]. As shown in fig 1.3, an IoT device is composed of five main units.

- **Processing unit:** provides all the computation tasks and manages the sensing, actuating, transmission and every other process. The diverse electronics of this unit mainly depends on the intended application. Different types of processing unit exist: micro-controllers, field programmable gate arrays (FPGA), system on chip

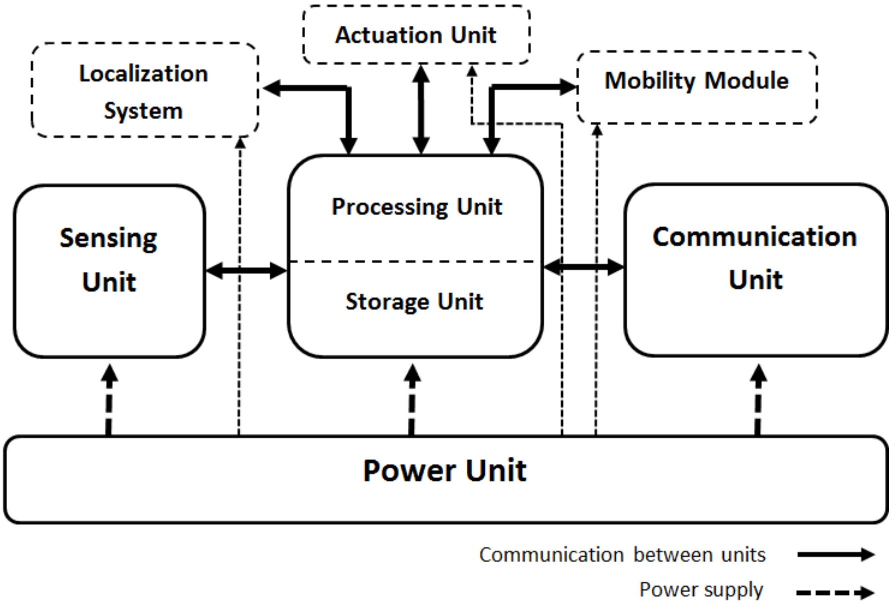


Figure 1.3: Typical smart device component

(SoC), multicore, complex programmable logic devices (CPLD), etc. [38].

- **Storage unit:** Usually consists of three components, represented by the RAM/SRAM (static random-access memory), Flash memory and EEPROM.

- **Communication unit:** allows the intercommunication and data exchange. It represents a high power-consuming module. Therefore, the choice, design and management of this unit is quite crucial especially in our context due to the energy constraint. Communications are generally Radio Frequency (RF) based.

- **Sensing unit:** represents the interface that enables nodes to collect data from the monitored environment and convert them into electrical signal to be analyzed by the processor. The electrical signal is digitized using Analog to Digital Converters (ADC).

- **Actuator unit:** allows automatic reaction of the device regarding the environment change. The actuators can act on their immediate surroundings to allow the proper operation of the machines or devices in which they are integrated. They can be divided into four main categories according to their construction model and the task they perform in a specific IoT context [39]: linear actuators, motors, relays and solenoids.

- **Power unit:** feeds all the device components with electrical energy. The power unit represents the major system constraint and determines the device lifespan. Therefore, the extension of battery life and is vital for the system efficiency.

1.4/ IoT and Wireless sensors networks (WSN's)

Wireless sensor network (WSN) is one of the emanations of IoT systems. WSN represents an ad hoc network composed of a large number of cooperative sensors wirelessly interconnected. Commonly, these heterogeneous devices work together to accomplish a particular goal, e.g. monitoring a particular large area. The sensor nodes are considered as source nodes. In order to forward the collected data to the external unit, some nodes act as gateways, also called sink nodes. Sink nodes are mostly connected to an external network to send the data to the final service.

WSNs provides the enabling technological foundations of IoT networks [40] including the intra-nodes communications such as Low Power Bluetooth [41] and ZigBee [42]. The complement functionality addressed by the IoT systems is the ability to interact with Internet [43]. To that end, IoT networks involve many telecommunication technologies such as LORA [44], IEEE 802.11 [45], etc. Therefore, the efficiency of IoT networks depends on the adaptation ability to nodes mobility, the optimisation of wireless technology usage and the energy consumption management. An overview of the wireless standards and technologies are provided in section 1.7

1.5/ IoT elements

Understanding the element composing the IoT helps to gain a better perception of the meaning and functionalities of this paradigm. In this section, we discuss the main elements involved in IoT systems as shown in Figure 1.4. Table 1.2 categories these elements and gives some examples of each category.

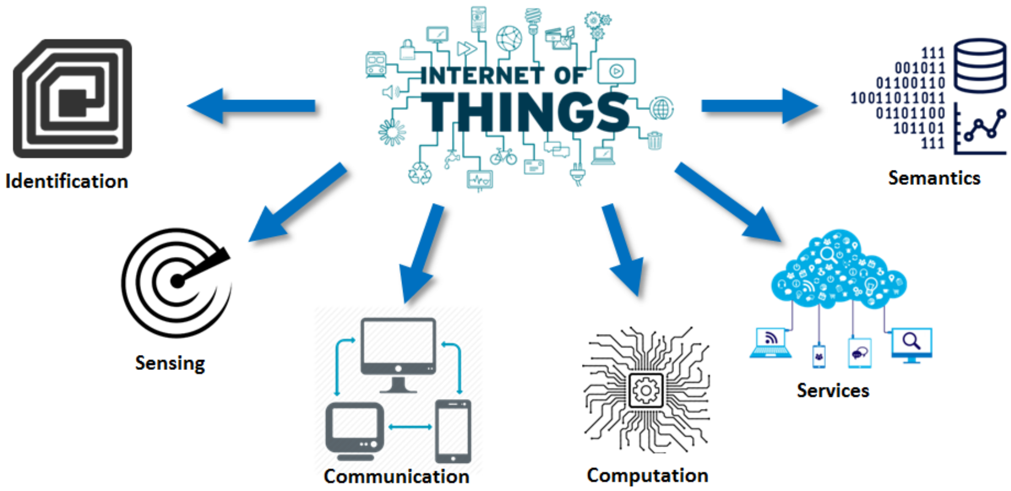


Figure 1.4: Basic IoT elements

1.5.1/ Identification

Identification is an essential task in IoT systems in order to name and match services with their client nodes. Many identification methods were developed for this purpose, such as electronic product codes (EPC) and ubiquitous codes (uCode) [15]. On the other hand, addressing the IoT objects is also crucial to distinguish between objects ID's and their addresses. The object ID refers to its name (e.g. "$S01$" for a particular sensor) and object's address refers to its identity within the network.

The addressing methods of IoT objects include IPv4, IPv6 and 6LoWPAN [46]. The 6LoWPAN eliminates many IPv6 overheads, allowing a tiny IPv6 datagram to be delivered over a single IEEE 802.15.4 hop in the best case scenario. It can compress the IPv6 headers up to two bytes [15].

1.5.2/ Sensing

The IoT sensing means gathering the environmental data by the network nodes and relaying them to a data warehouse, a database, or a cloud. The collected data are analyzed to take particular actions.

1.5.3/ Communication

The IoT communication technologies connect heterogeneous objects together in order to exchange data. Commonly, the IoT devices operate using low power protocols under noisy communication links. Several communication protocols are used for the IoT, such as WiFi, Bluetooth, IEEE 802.15.4 (LR-WPAN), Z-wave, and LTE-Advanced. Some specific communication protocols are also used like NFC (Near Field Communication), UWB (ultra-wide bandwidth). RFID technology represented one of the first protocols used to achieve M2M communication (RFID tag and reader) [15].

1.5.4/ Computation

The computational ability of the IoT are supported by the processing units (e.g., microcontrollers, microprocessors, etc) and the software applications. Diverse hardware platforms were designed to run the IoT applications, such as, Arduino, UDOO, FriendlyARM, Intel Galileo, Raspberry PI, Gadgeteer, BeagleBone, Cubieboard, Z1, WiSense, Mulle, and T-Mote Sky [15]. Accordingly, various operational systems were developed to manage the functionalities of the devices. These Real-Time Operating Systems (RTOS), such as Contiki RTOS, TinyOS, LiteOS and RiotOS, offer light weight OS intended for IoT environments [15]. In Table 1.1, certain properties of these operating systems are compared.

Table 1.1: Common RTOS based operating systems in IoT [15]

Operating System	Supported language	Required Memory	Event-based Programming	Multi Threading	Dynamic Memory
Contiki	C	2 KB	Yes	Yes	Yes
TinyOS	nesC	1 kb	Yes	Partial	Yes
LiteOS	C	4 kb	Yes	Yes	Yes
RiotOS	C/C++	1.5 kb	No	Yes	Yes
Android	Java	32 Mb	Yes	Yes	Yes

1.5.5/ Services

IoT services can be categorized under four classes [15]. Firstly, the Identity-related services which represent the most basic and essential services that are used in other types of services. Every application that brings real world objects to the virtual world needs to identify and authenticate those objects. Second category includes Information Aggregation Services that gather and resume sensors' measurements in order to be processed. Thirdly, the Collaborative-Aware Services that act on the Information Aggregation Services and use the collected data to make decisions and behave accordingly. Finally, Ubiquitous Services that ensure Collaborative-Aware Services anytime, to anyone, and anywhere. For instance, the majority of the current applications provide identity-related, information aggregation, and collaborative-aware services. Smart healthcare and smart grids belong to the information aggregation category, wherease, smart home, smart buildings, intelligent systems (ITS), and industrial automation are closer to the collaborative-aware category. The ubiquitous services is an objective for all IoT applications. However, reaching this goal is not easy since there are many difficulties and challenges that need to be addressed.

1.5.6/ Semantics

This element refers to the ability of the service to extract relevant knowledge from the IoT devices in order to render the required services. This extraction covers the recognizing and analyzing of data to realize the right decision or provide the best service. This requirement is assisted by Semantic Web technologies like Resource Description Framework (RDF), Web Ontology Language (OWL), and XML Interchange (EXI) [15].

categories these elements and gives some examples of each category.

Table 1.2: IoT elements and technologies

IoT Elements		Examples
Identification	Naming	EPC, uCode
	Addressing	IPv4, IPv6, 6LoWPAN
Sensing		Sensing devices, Smart sensors, Wearable, Actuators, RFID
Communication		RFID, NFC, UWB, Bluetooth, BLE, WIFI, Z-Wave, LTE-A, IEEE 802.15.4 (LR-WPAN)
Computation	Hardware	Arduino, UDOO, FriendlyARM, Intel Galileo, Raspberry PI, Gadgeteer, BeagleBone, Cubieboard, Z1, WiSense, Mulle, T-Mote Sky
	Software	Contiki RTOS, TinyOS, LiteOS and RiotOS
Services		Identity-related (e.g. shipping), nformation Aggregation (smart greed), Collaborative-Aware (smart home), Ubiquitous (smart city)
Semantics		RFD, OWL, EXI

1.6/ IOT NETWORK TOPOLOGY

The conception of an operational wireless network in a real world environment is a cumbersome task which can be further complicated due to many unanticipated scenarios, including the dynamicity of the environment, topography, the wireless channel and many other factors that cannot be modeled with exactitude beforehand. The network topology represents the physical or logical layout of the network. The random deployment of network devices produces different shapes. An appropriate topology is able to ensure an adequate connectivity of the network and facilitate the wireless communications. There are two main types of wireless network topology, namely the flat and the hierarchical topology.

1.6.1/ THE FLAT TOPOLOGY

The flat topology is the simplest in terms of structure, all network devices, except for the sink gateway, have the same task of collect and send their contextual data. A sensor can directly communicate with the sink either in a single-hop (e.g. star network) or in multi-hop mode. The simplicity of this architecture enables low communication latency. However, the distance between the furthest node and the sink is limited because transmission costs increase in terms of energy consumption, and in the worst-case scenario, the base station may become unreachable. This scenario may be prevented by using multi-hops transmissions. However, when the network density increases, collisions and interference's issues arise, because the wireless medium is shared and maintained by individual nodes in the flat model, which results in low efficiency in the resource usage. Consequently, this topology is often applied in indoor scenarios (e.g. offices or courtyards) and are not suitable for large scale scenarios such as environmental monitoring.

1.6.2/ The Hierarchical topology

The hierarchical architectures are typically used in large wireless networks. It is composed of three types of nodes: the sink, routing nodes and sensor nodes. Network nodes are organized into clusters or a tree where the sink plays the role of a root. Sensor nodes transmit their data to their own routing parent node. Routing nodes play the role of route planners, they execute routing algorithms to find the suitable path toward the sink. In a cluster based topology, the network is divided into groups or clusters with a designated node in each cluster that acts as a Cluster Head (CH). The CH coordinates the cluster nodes and aggregates the received data before forwarding them to the sink. Its worth mention that the network nodes perform the routing and the CH tasks in addition to their basic functionality (typically sensing). Therefore, in most cases, these tasks are undertaken temporarily by different nodes to load balance the energy consumption among network nodes. Although, the hierarchical topologies may increase the communication latency and faces the drawbacks of structure complexity and energy efficiency due to route planning. These structures overcome the limits of flat topology and therefore are more appropriate for large scale applications. Figure 1.5 highlights a classification example of network topologies with (a) Single-hop flat model, (b) Single-hop clustering model, (c) Multi-hop flat model, (d) Multi-hop clustering model.

1.7/ Wireless communication standards

IoT network is based on several and heterogeneous wireless technologies. Communication protocols define how the radio resources are shared by the different devices in order to guarantee the efficiency (throughput, capacity), the reliability (blocking rate, interference, etc.) and durability (energy). A communication protocol allows network devices to send and receive information packets and define different transmission aspect, such as, the data exchange formats, the data encoding, addressing schemes, routing of packets, flow control, etc. These Wireless access technologies can be classified into three main categories, namely, WLAN (Wireless Local Area Network), cellular and satellite technolo-

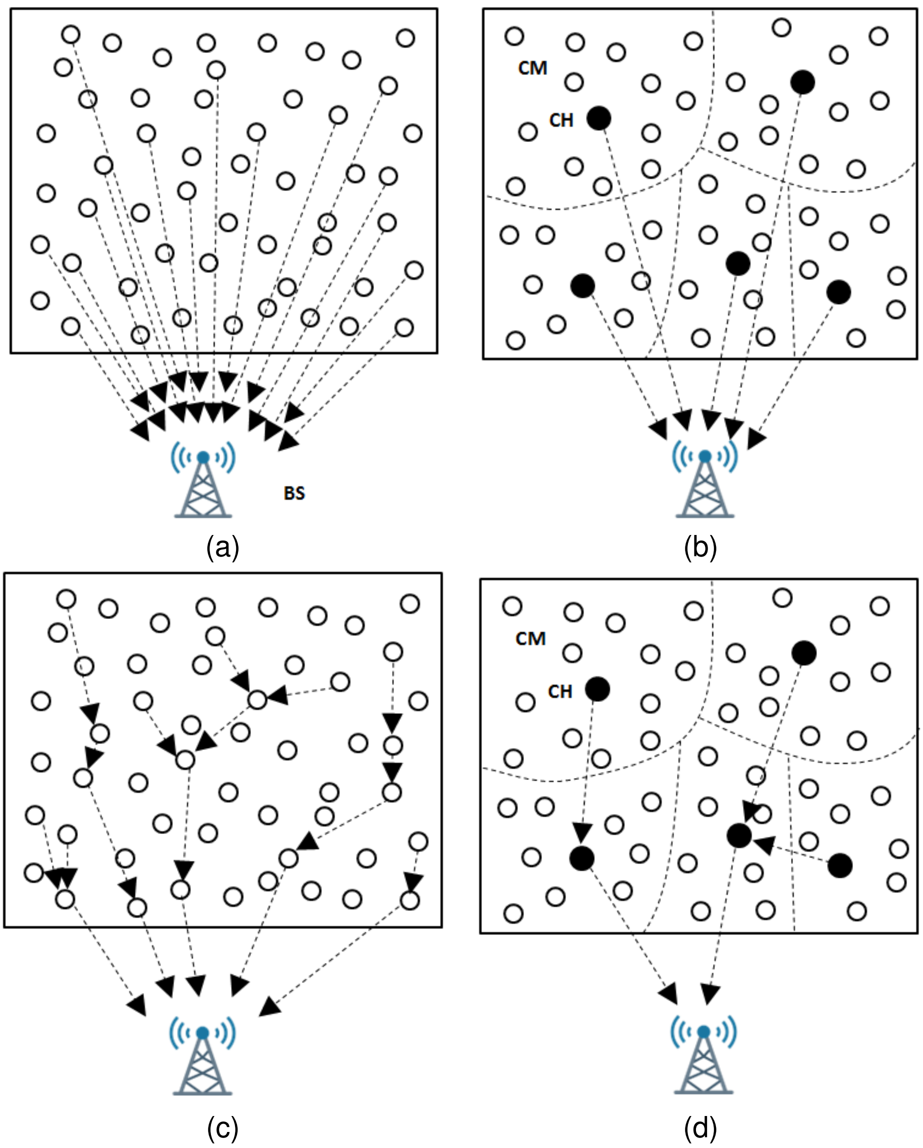


Figure 1.5: Classification of network topologies: (a) Single-hop flat model, (b) Single-hop clustering model, (c) Multi-hop flat model, (d) Multi-hop clustering model

gies as shown in Figure 1.6.

1.7.1/ Wireless Local Area Network (WLAN)

WLAN networks have been gradually approved as a versatile connectivity solution for different modern Ad-hoc networks and environments. They are widely used due to their reduced cost, high data rates provided and their easy implementation. This category regroups many standards.

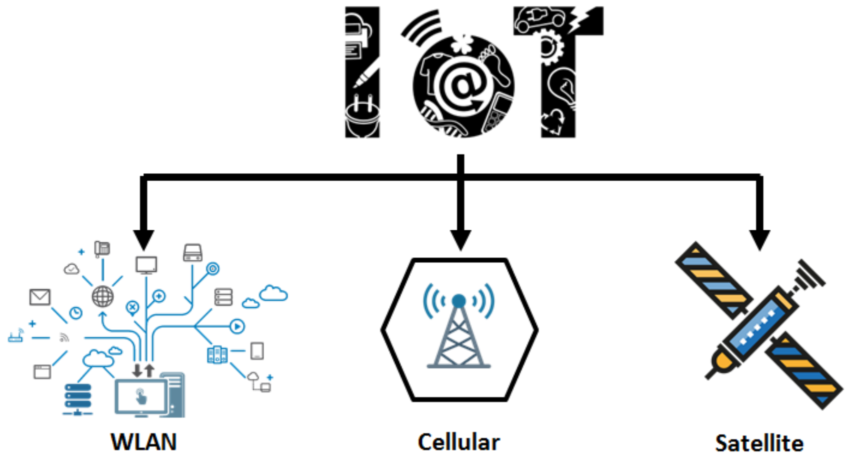


Figure 1.6: Wireless access technologies in IoT

1.7.1.1/ IEEE 802.15.4 - LR-WPAN

The Low-Rate Wireless Personal Area Networks (LR-WPAN) are represented by the IEEE 802.15.4 which constitutes a collection of standards that form the basis specifications of high-level communication protocols. The IEEE 802.15.4 standard is adopted by many low-cost wireless interfaces such as Zigbee, WirelessHART [47], 6LoWPAN [46], Z-Wave [48], etc. In addition, advanced communication technologies for low-power long-range were proposed in the literature to manage even wider communication areas [15], such as BLE (Bluetooth Low Energy) [41] which overcomes the Bluetooth in terms of power consumption while maintaining a similar transmitting range.

LR-WPAN protocols are low-cost and provide low-speed communication to power constrained devices. These standards offer a data rates between 40 Kb/s and 250 Kb/s and operate at 868/915 MHz and 2.4 GHz for low and high data rates, respectively.

1.7.1.2/ IEEE 802.15.1 - BLUETOOTH

The Bluetooth is a low power and low cost wireless communication technology suited for short range transmissions (lower than 10m). The Bluetooth standard defines a Personal Area Network (PAN) communication based on the IEEE 802.15.1 standard. It operates under the 2.4 GHz frequency and can provide a data rate up to 3 Mb/s [15].

1.7.1.3/ IEEE 802.11 – WIFI

The IEEE 802.11 (WIFI) reached a great acceptance in domestic, academic and industrial domains, it is characterized by a short/mid-range supports and a relatively high-speed data transmission. The IEEE 802.11 standard specifies the set of medium access control (MAC) and physical layer (PHY) protocols in order to implement a WLAN. The WIFI represents a collection of WLAN communication standards (e.g. IEEE 802.11a/b/g/n/p/ac), it provides an approximate communication range of 20 m indoor and 100 m outdoor. IEEE 802.11 standards provide data rates from 1 Mb/s to 6.75 Gb/s. For instance, the 802.11a operates with the 5 GHz band, 802.11b and 802.11g operate in the 2.4 GHz frequency, the 802.11n operates with 2.4/5 GHz bands and 802.11ac operates in 5 GHz [45].

1.7.1.4/ IEEE 802.16 – WiMax

WiMAX (Worldwide Interoperability for Microwave Access) is a collection of wireless broadband communication standard. It is based on the IEEE 802.16 set of standards and it provides multiple physical layer (PHY) and Media Access Control (MAC) options. IEEE 802.16 provides data rates from 1.5 Mb/s to 1 Gb/s. For example, the 802.16m (WirelessMAN-Advanced) reaches a data rate of 100 Mb/s with mobile stations and can reach up to 1 Gb/s for fixed stations. This latter was a candidate in competition with the LTE Advanced standard for 4G mobile networks.

1.7.1.5/ LoRaWAN

LoRaWAN (Long Range Wide-area Network) is a recently developed protocol (2015) to maintains long-range transmission even in rural areas with low power consumption. Lora standard provides a data rate in the range of 0.3 kb/s to 50 kb/s and operates in 868 and 900 MHz, it can reach a transmitting range of 15 km in unobstructed environments. LoRaWAN was designed by the LoRa Alliance [6]. LoRaWAN is a best candidates for WSN applications that are insensitive to delay and do not need high data rates.

WAVENIS [49] characterized by its ultra-low power and long-range communication, used in building management, chemical and healthcare applications. Table 1.3 shows a brief comparison of some prevalent WLAN's technologies with their associate application.

1.7.2/ Cellular Wireless technology

Cellular Wireless technologies have been widely used to provide mobile services since the 1970s. They mainly use radio waves to transmit data over long distances with a frequency reuse technique (except in UMTS) to increase the network coverage and allow

[6] LoRa Alliance: a nonprofit association created in 2015 by Cycleo https://lora-alliance.org

Table 1.3: Comparison between presented clustering algorithms properties

Protocol	Frequency Band	Data Rate	Range	Energy Consumption	Usage Area
Wi-Fi [45]	2.4/5.0 GHz	54 Mb/s	30 m	Medium	Home entertainment, Industrial
Bluetooth [15]	2.4 GHz	1 Mb/s	10 m	Low	Industrial, Traffic management
IEEE 802.15.4 [42]	868–915MHz /2.4GHz	250 kb/s	[10, 100] m	Low	Industrial, Traffic management, Smart home, vehicle monitoring
LTE [50]	2.6 GHz	10 Mb/s	≤15 km	High	Mobil telecommunications, Smart Cities
Wireless HART [47]	2.4–2.5 MHz	250 kb/s	[1, 100] m	Low	Healthcare, Industry automation
BLE [41]	2.4 MHz	1 Mb/s	200 m	Very Low	Healthcare, Home entertainment
Z-WAVE [48]	1 GHz	40 kb/s	30 m	Low	Home automation applications
WAVENIS [49]	865–916 GHz	100 kb/s	≤4 km	Very Low	Chemical and healthcare applications
LoRaWAN [44]	868–900 GHz	50 kb/s	≤15 km	Very Low	Smart City, industrial Monitoring, Agriculture
NB-IoT [51]	180 kHz	234.7 kb/s	≤35 km	Low	Industrial Monitoring, Smart City

simultaneous transmissions. In this kind of networks, the covered area is divided into small cells each covered by a base station that serve the nearest users. Traditional cellular network use the frequency band between 450mhz and 3 GHz [24]. All the BSs are connected through a fast wired or wireless links. These networks are typically modeled using hexagonal grids to represent the cells with a BS placed at the center of the cell as shown in figure 1.7. However, with the alteration of user demands across the service and the impossibility of the BS in harsh environment (e.g. rivers, rails, hills, buildings, etc.), a special randomness in the BS placement and the coverage area of a cell may vary from a region to another. Cellular wireless technology comprises different generations of mobile communication standards.

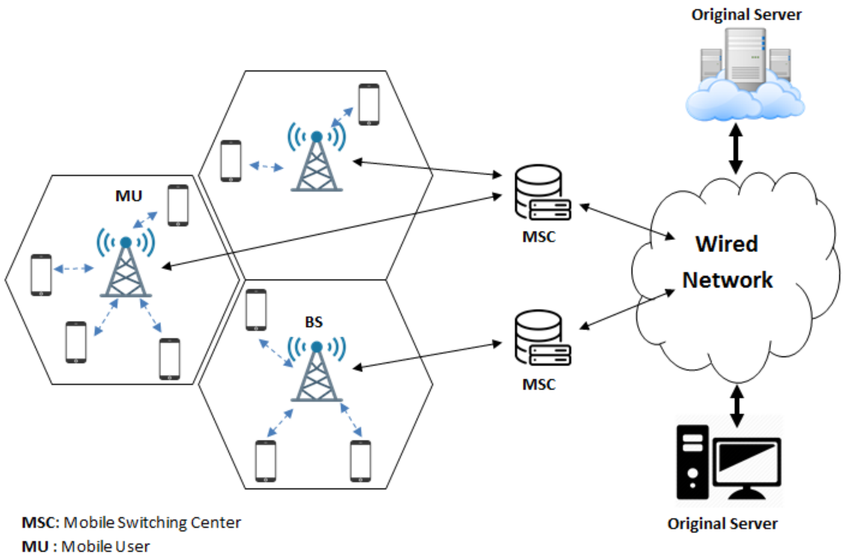


Figure 1.7: Wireless cellular network architecture

The 2G (Second Generation) enables better mobile voice communication over 1G, the Global System for Mobile (GSM) was integrated with a frequency band between [890-915 MHz] and the maximum achievable speed was 1 Mpbs [52]. The third generation 3G (UTMS) higher system speed and quicker download. The technology behind 3G was high-speed packet access (HSPA/HSPA+). The 3G employed MIMO (Multiple Input Multiple Output) technology for multiplying the power of the wireless network, it also used packet switching for quick data transmission. The 3G is able to achieve a data rate of 384 Kbps with moving devices and go up to 2 Mbps with stationary nodes [52]. The fourth generation 4G (LTE) is designed to support a high mobility and offers a higher speed and broadband with a fast mobile internet access. It works on LTE and WiMAX technologies, as well as provides wider bandwidth up to 100 Mhz. The Fourth Generation LTE-A (4.5G), an advanced version of standard 4G LTE, uses MIMO technology to combine multiple antennas for both transmitters as well as a receiver. This technology makes the LTE-A three times faster than standard 4G. LTE-A delivers speeds of over 42 Mbps and up to 90

Mbps [52].

The fifth generation 5G begins deploying in 2019 and provide connectivity to the most current cellphones. 5G networks are predicted to reach more than 1.7 billion subscribers by 2025. These modern networks will make possible new applications of the IoT paradigm. It provides down-link maximum throughput of up to 20 Gbps [16]. To support the 5th Generation (5G) communication paradigm, various wireless technologies have evolved, such as LoRaWAN, Narrow-Band IoT (NB-IoT) and Long-Term Evolution for Machines (LTE-M) [16].

NB-IoT is launched by the Third Generation Partnership Project (3GPP). The object of NB-IoT is to scale the network devices to be usable and more dependable. The constant synchronization needed in NB-IoT requires more energy (compared to LoRaWAN), besides that more peak current is needed in Orthogonal Frequency-Division Multiplexing (OFDM) or Frequency Division Multiple Access (FDMA) for the linear and constant transmission. Moreover, the battery life time of NB- IoT node is shorter than that in LoRa nodes. Therefore, NB-IoT is suited for applications that need higher data rates transmission and low latency [16].

LTE-M stands for Long Term Evolution for Machines. It is a new narrowband Machine to Machine (M2M) technology based on the existing LTE network. M2M refers to the communication model that allows nodes to directly communicate with each other. LTE-M has been introduced to be the major communication technology for future urban systems. The important specifications of this new technology are the low cost, especially on the node side, high coverage, the long battery life, and the high capacity (e.g. a larger number of nodes per cell). The low-cost feature due to that the bandwidth requires less expensive radio frequency units. The coverage is enhanced due to the ability to focus the transmission power in narrow bandwidth [16]. Table 1.4 shows a comparison between the 5G emerging communication technologies (e.g. LoRaWAN, NB-IoT, LTE-M)

1.7.3/ Satellite communications

The Satellite communication technology is based on the usage of the available satellite system in the communication field. It offers various mobile services, e.g. voice, radio channels, video calling, Internet access and television. Additionally, this technology can provide ubiquitous coverage and an extension of communication capabilities over long distances. Satellite communication represents an appropriate solution to work under inoperable communication conditions, such as: desert, mountain, oceans [53]. The drawbacks of this technology are their expensive cost and the large propagation latency, which affects negatively the network performance. Table 1.5 illustrates a comparison of long range wireless access technologies.

Table 1.4: Emerging communication technologies in 5G [16]

Access Technology	Modulation	Cost	Channel band-width	Data rate	Coverage	Durability (2000mAH)
LoRaWAN	CSS	Low	125-500 KHz	50Kbps (UL)/290 Bps	5Km (urban)-20Km (rural)	105 months
NB-IoT	GFSK (UL)/BPSK(DL)	Medium	180 KHz	220 Kbps	11Km (ur-ban) -10Km (rural)	90 months
LTE-M	SC-FDMA, (UL) / OFDMA, (DL)	High	1.2 - 20 MHz	1 Mbps	5 Km	18 months

Table 1.5: Long range wireless access technologies in IoT

Access Technology	Standard	Frequency	Coverage	Interference
WLAN's	802.11 a/b/g/n/p/e/m	[2.4-5.9] GHz	$\leq$50 km	Medium
Cellular	UMTS/LTE	450MHz - 5.9 GHz	50 km	Low
Satellite	MSIA/DVB	30 GHz	2000 km	Low

1.8/ NETWORK REFERENCE MODEL

As a network system, the IoT network obeys to layered architecture model. The networking reference model is an organizational architecture used to design the network. It allows a better comprehension of the functioning of the network. The system communication task is partitioned into multiple sub-tasks and the related sub-tasks are mapped into layers. The different tasks of a layer are implemented by protocols. The usage of a layered reference model for communication system enables an easy integration of new services within the dedicated sub-tasks layer.

Several reference models have been proposed in the literature by different researchers [20; 15]. These models are inspired from the most popular called Open System Interconnection (OSI) model with 7 layers. The TCP/IP model with 4 layers is a conventional model widely used in the Internet. However, this model may be inadequate for the IoT for the following reasons [17]:

- *Connection setup*: the TCP is a connection-oriented protocol, and each session begins with a connection establishment procedure (the three-way handshake). This is unnecessary because the majority of IoT conversations will entail the transmission of tiny amounts of data. Hence, the setup phase would take up a significant percentage of the session time.

- *Congestion control*: TCP is in responsible for carrying out end-to-end congestion management. This may generate performance issues in the IoT because most connections will take place over the wireless medium, which is known to be a difficult environment for TCP [71]. Furthermore, if the quantity of data to be transferred in a single session is extremely tiny, TCP congestion management is pointless, because the whole TCP session will be closed with the transmission of the first segment and the subsequent receipt of the appropriate acknowledgement.

In addition, some protocols, such as IPv6, may be inapplicable with the low power Wireless Personal Area Network (LR-WPAN), which is commonly used in IoT, since low power WPANs have low bandwidth, tiny packet sizes, and battery powered devices with restricted processing capabilities. Consequently, there was a requirement to create an adaptation layer that conformed IPv6 packets to LR-WPAN (IEEE 802.15.4) requirements. In 2007, the IETF 6LoWPAN working group created such a standard, called 6LowPAN [46]. This group work on specifications that would allow IPv6 packets to be carried via IEEE 802.15.4 MAC. They characterize the encapsulation and the header compression to decrease the overhead, as well as the fragmentation to meet the IPv6 Maximum Transmission Unit (MTU) requirement.

Therefore, due to the explicit nature of the wireless networks, a five layer model (figure 1.8) is being adequate to model the functional division of the wireless network. Especially for WSNs which constitute a fundamental part of IoT networks [13].

In this latter, the application layer regroups the functions of the application, presentation and session layers present in the OSI model. Whereas, the functions of the transport, network, data link and the physical layers are similar to the OSI model. In the following, we provide a brief description of the functions of these layers.

- **Application layer** interacts with the human user and makes the lower layers transparent. Its functions include application processing, external database access, queries processing, data encryption, authentication, etc.

- **Transport layer** guarantees the managing of end-to-end communications, it ensures the transfer of data packets without corruption, loss or duplication. Sometimes transport layer just guarantees the detection of transmission errors or the congestion control.

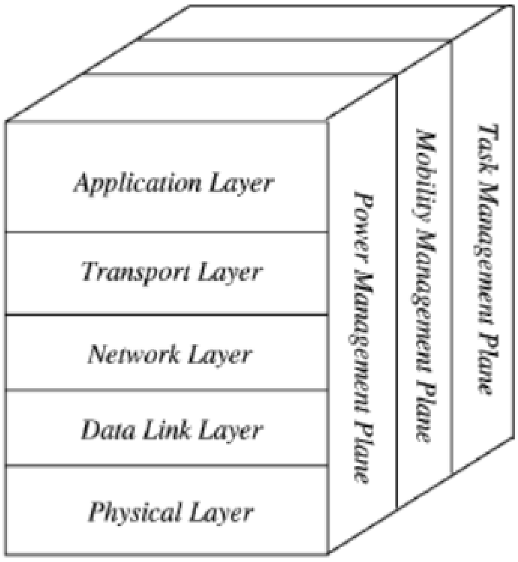


Figure 1.8: Wireless network reference model [13]

- **Network layer** mainly manages the routing of data packets in the network. The routing policy may include an adaptive topology management.
- **Data Link layer** is responsible of the recognition of the data frames, and controls the access to the radio channel. It is constituted by 2 sub-layers:
  - **Media Access Control (MAC) sub-layer**: the function of this sub-layer includes the channel sharing, admission control (new communication acceptance) and the detection of transmission errors.
  - **Logical Link Control (LLC) sub-layer**: controls the flow of data.
- **Physical Layer** offer an interface for transmitting a flow of bits stream via a physical medium. It is responsible for for ensuring the signal detection, carrier frequency generation, data encryption and modulation.

Beside this five layers (inspired by the OSI model), there are three cross layers management plans used to provide more functionalities and increase the overall efficiency of the network. These include power, mobility and task management planes, as shown in figure 1.8.

- **Power management plane:** manages the energy consumption of the device. For example, in case of an energy critical level, the concerned device can inform its neighbors that he can no longer participate to the routing task.
- **Mobility management plane:** detects the device movement and monitors the neighbor's movements and power levels (for power balancing purposes).

- **Task management plane:** is responsible of the tasks management that ought to be performed by the device.

1.9/ IoT Application

Due to the capabilities offered by the IoT networks, diverse new applications have emerged, among which only a small part is currently available to our society. This novel infrastructure builds the basic network required by many valuable applications that need a consistent connectivity and addressability between the network components. There are many domain and environment in which IoT applications would likely improve the quality of our lives (e.g. living home, work location, transport, agriculture, environment monitoring, entertainment, public health, etc.). These environments are mostly equipped with primitive intelligence objects, giving these devices the possibility to communicates with each other and elaborate the information perceived from their surroundings environment will deploy a very wide range of applications. Among these potential applications, we may distinguish between applications available today and those that could be qualified by futuristic applications, which are only fancy for the moment (due to the limits of the available technologies) [17]. In the following we provide an overview of these applications (figure 1.9).

1.9.1/ Smart home

IoT services attempt to improve the user's lifestyle by making it easier and more appropriate to monitor and remotely handle home appliances and systems (e.g., air conditioning, heating and energy monitoring systems, refrigerators, laundry machines, lighting, etc.). Sensors and actuators deployed around the house can make our life more pleasant in several aspects. For example, a smart house can automatically adapt the rooms heating according to the weather and the resident's preferences, adapts the house lighting according to the daytime or reducing the energy consumption by automatically switching off the unused electrical equipment. Moreover, the energy consumption may be optimized by considering when the electricity prices are cheap (information provided by an external web service and according to the current energy production) and by considering the specific requirements of each house appliance. The interest for such houses lies in the fact that they permit a remote control of house devices and allow an efficient management of the energy usage [15].

1.9.2/ Smart factory

A smart factory is a production facility made up by a number of sensors and robotic connected devices that intend to enhance the industrial process through automation and self-optimization (with a minimal human involvement). It allows a quick and accurate production based on four components: processing, transportation, sensing and communication. The IoT is used in the industrial field to control and monitor production machines functionalities and their productivity through the Internet. This paradigm will increases the industrial productivity by analyzing the production data, timing and causes of potential problem. For example, in case of a production line issue, the system can automatically report the issue to the maintenance service to recover the problem. Moreover, a wireless sensor mounted on the machine may automatically monitor the machines vibration in order to control the process and preserve the product quality. Consequently, Smart factories promote the productivity and reduce the cost of production [17].

1.9.3/ Smart gym

The gym is a representative example where the IoT paradigm can allow a better exploitation of the gym facility. For example, the instructor can upload the exercise profile to the training machine for each trainee, which is recognized by the machine through his RFID tag. The health parameters of each individual is controlled during the whole session and reported to check whatever the trainee is over-training or is relaxed when doing the exercises [17].

1.9.4/ Smart cities

A smart city could be seen as a city that uses several ubiquitous services and different devices such as sensors or actuators to collect and analyze the data (e.g. from the residents or from the environment), then take a decision that aims to improve the life quality in the city. In such environment, various smart systems are interconnected to offer the required services (health, transportation, homes and buildings assistance, etc.). The benefits of a smart city can be observed in many scenarios, notably the energy conservation system, water supply distribution, waste management, traffic and parking management, environment monitoring, etc [15; 17].

1.9.5/ Smart healthcare

Smart healthcare plays a significant role in healthcare applications; it corresponds to health system where sensors and wearable devices are used to improve the diagnosis and the treatment of the patients. This system will allow a real-time monitoring of the physiological statuses of patients through sensors by analyzing the collected information remotely (by the concerned doctor) in order to proceed for the treatment. Such a system improves the efficiency of clinical care by a quickly access to the patient information, which in turn will be satisfied by the outcomes of their early treatments. As an example of medical wearable devices, Masimo Radical-7 remotely monitors the patient's status and send the patient medical reports to the clinical staff [7]. Another example is the prevention from the medical infections. In this field, IBM scientists along with Ohio Health's hospitals utilized the RFID technology and real-time Big Data analysis to track hand washing of medical staff after checking each patient. That collaboration was introduced to avoid infections that cause about 90 000 deaths in the united states and a loss of about 30 billion dollars annually [54].

1.9.6/ Smart transportation

A Smart transportation system or intelligent transportation system (ITS) refers to a transport system based on modern technologies and management strategies. It represents an integration between computation and communication to supervise and control the transportation network. These systems employ several components, such as the cars subsystem (GPS, RFID reader and the communication system), the roadside equipment, and the monitoring center and security subsystem. The ITS aims to make driving more reliable, enjoyable and efficient. Moreover, this system can be used in vehicles for collision avoidance (rode safety purpose), help to reduce traffic congestion and improve the quality of the environment. For instance, AUDI is the first automaker with a license for self-driving in Nevada, USA. Volvo announced in 2013 it first self-driving car to drive about 30 miles in busy roads in Gothenburg, Sweden [15].

1.10/ Requirements and challenges of IoT networks

1.10.1/ Requirements

The ultimate objective of IoT applications is undoubtedly to achieve the ubiquity of services. It is noted that the IoT paradigm is not complicated, however, realizing this vision is not an easy task due to the requirements that need to be addressed. Several applications with shared actuator devices and sensors need to operate simultaneously to form a reliable system. With a hyper-connected IoT network, a design error in a part of the system may lead to a system failure, which may be chaotic. Therefore, there are

[7]Masimo Radical-7: A CO-oximeter connected device that measures the oxygen carrying state of hemoglobin in a blood specimen https://www.masimo.fr/products/continuous/radical-7/

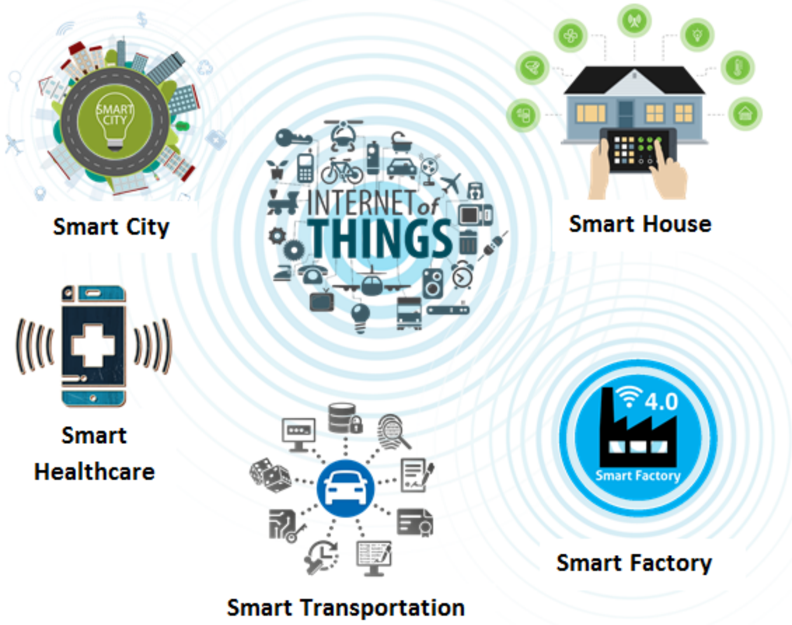


Figure 1.9: Overview of IoT networks applications

some key requirements that need to be addressed, including scalability, energy efficiency, fault tolerance, quality of service, availability, reliability, interoperability, security and management, etc. Most of the identified requirements and challenges are stated in surveys [15; 17; 55; 56; 22]. Some research projects, e.g. IoT6 and RELYonIT [8] investigate the challenges and weakness of the IoT paradigm and attempt to provide suggestions and guidance for the solutions [15]. In the following, we briefly discuss the key challenges that face the development of the IoT network. Table 1.6 exhibits the research works that render the common IoT challenges and the requirements that ought to be considered during the design of the IoT architecture.

1.10.1.1/ Scalability

IoT networks are characterized by the presence of a massive number of resource constrained devices. The scalability refers to the ability of adding new devices and services without affecting the quality of the existing network. Due to the presence of diverse hardware platforms and communication protocols, adding new system operations may be difficult. Therefore, new proposed approaches need to satisfy the scalability requirement in order to avoid poor network performances.

[8] http://www.relyonit.eu/

1.10.1.2/ Energy efficiency

One of the main requirements of IoT networks is the energy efficiency. The reduction of energy consumption is a critical concern in wireless networks due to the energy constraint of the devices. An energy efficient network has the ability to execute operations with a minimum amount of energy in order to maximize the network lifetime. This requirement is especially desirable in IoT networks as they are typically composed of devices with limited batteries. In addition, devices death may lead to a loss of connectivity which may cause an entire network interruption.

1.10.1.3/ Fault tolerance

Fault tolerance is the property that enables a system to keep operating properly even with the failure of some elements and mitigates the component failure effects [57]. This requirement is crucial to ensure that the network is able to fulfill its expected functioning without interruption even in case of faults or crashes (e.g. devices issue, network fault, sink fault, software problem, etc.). This property is necessary for IoT networks as they may be liable to devices failure because of their characteristics (energy, processing, and storage) or their environment (e.g. pipeline, chemical spill area, etc.).

1.10.1.4/ Availability

The achievement of an available IoT service requires the achievement of the availability of hardware and software. As the IoT devices are prone to failure [37], the main challenge that face the design of an available system is the tolerance to failures as they occur and recover from their effects. Therefore, fault tolerance methods and the redundancy of component and services is required. With the redundancy, the availability of the hardware can ensure the application availability [15]. For instance, in case of a component crush, the service can be maintained by another component. However, the problem of using the redundancy is the cost of the generated application.

1.10.1.5/ Quality of services (QoS)

The dynamicity of the network and the potential transient instability can cause frequent disconnections in the communication links which may temporary affect the performance of the IoT application. The quality of services (QoS) refers to the evaluation and the optimization of overall service performance and user satisfaction. This performance is evaluated using different metrics, such as, packet loss, communication latency, data throughput, bandwidth and energy consumption, etc. Therefore, the QoS required in the network depends on the targeted application. For example, applications like smart lighting are delay tolerant while other IoT critical applications (e.g. fire detection) are critical and require a short aggregation delay (alert notification).

1.10.1.6/ SECURITY

Security is an important concern for IoT network as reviewed in [58; 59]. According to [37], 70% of the IoT devices are vulnerable and are subjects to malwares. This problem is caused by the lack of transport encryption, inadequate software protection, and insufficient authentication due to the limited capacities of the network devices. In addition, security attacks can easily affect the network for many other reasons, notably, the wireless communication between devices, the openness of the system, or the physical accessibility of the devices. A particular attention should be paid while designing secure network protocols because the security mechanisms developed for traditional network may not be convenient for resource-constrained devices.

1.10.1.7/ RELIABILITY

The reliability refers to the ability of the system to work properly according to its specifications [60]. The reliability intends to enhance the success rate of IoT delivery service and must be implemented in all the networking layers [15]. This criterion is strongly related with the availability, i.e. the availability of a service can be reached over time by the insurance of the reliability [15]. The Reliability is a stringent requirement especially in the field of critical applications [61]. In these systems, the communication network is an important component that needs to be fault resilient in order to permit a reliable information distribution and prevent errors in data gathering, data processing,decision, and transmission.

1.10.1.8/ MANAGEMENT

Connecting a large number of objects to the network makes the management of the faults, configurations, accounting, performance and Security a difficult task. To deal with such a problem, the development of lightweight management protocols is required. This management is a key solution to assist the deployment of the IoT applications [15]. For example, monitoring Machine to Machine (M2M) communication of the IoT devices is essential to ensure all times connectivity and reach the users satisfaction. The Open Mobile

Alliance [9] developed a standard called Light-weight M2M (LWM2M) [10] to provide an interface between M2M devices and servers. LWM2M aims to build an agnostic scheme for the management of the diversity of devices [15].

1.10.1.9/ Interoperability

The interoperability is another requirement that face the deployment of the IoT network due to the large number of heterogeneous network devices that belong to different platforms. To satisfy the IoT customers (regardless of the used hardware or platform), the interoperability needs to be considered by both application developers and device manufactures. For example, to ensure the interoperability in nowadays scenarios, modern devices need to support the common communication technologies such as Bluetooth, WiFi, GSM, LTE, etc. Moreover, IoT programmer need to design application that allows an easy adding of new functionalities without issues, while maintaining the integration with different communication technologies. For instance, to avoid interoperability problem, the European Telecommunications Standards Institute (ETSI) [11] tests the interoperability between different implementations in order to assess the compatibility of the tested services. Research project like PROBE-IT [15] aims to ensure the interoperability of validated IoT solutions, such as CoAP and 6LoWPAN protocols, using interoperability tests [46].

1.10.2/ Challenges

Reaching the requirements of the IoT networks described previously remains an open issue and a research challenge due to the resources restricted devices, the dynamic nature of the networks and the constraints related to the deployment environment (e.g. war zone, pipeline, etc.). Different network management protocols have been proposed in the literature to meet those requirement (table 1.6). However, the proposed approaches lack when addressing the heterogeneity and the resource-constrained aspect of the devices. Hence, to fulfill the network requirement, the main challenges consist of developing mechanisms and protocols that ensure the satisfaction of all the network requirements with the consideration of the constrained aspect of the devices [33].

[9] Open Mobile Alliance (OMA): a standards organization which develops open, international technical standards for the mobile phone industry. https://omaspecworks.org/

[10] Light-weight M2M (LWM2M): a protocol from the Open Mobile Alliance for machine to machine (M2M) or Internet of things (IoT) device management and service enablement. (about LWM2M)

[11] ETSI: an independent organization that supports the development and testing of global technical standards for ICT-enabled systems, applications and services. https://www.etsi.org/

Table 1.6: Main research challenges addressed in IoT

IoT Challenges	Related researches
Scalability	[62; 15; 63]
Energy efficiency	[31; 30; 37; 36; 23; 33; 17]
fault tolerance	[57]
Availability	[37; 15; 64]
Quality of services	[65; 66]
Security	[58; 59; 56; 15; 65; 55]
Reliability	[15; 60; 61]
Management	[46; 25; 24; 35; 15]
Interoperability	[15; 62; 63]

1.11/ CONCLUSIONS

The Iot is an emerging paradigm that attracts expanding attention in both academic and industrial areas. IoT technology has allowed the emergence of effective applications in diverse domains. This paradigm aims to realize a smart world vision by facilitating our lifestyle. In this chapter, we provided an overview of the IoT network by highlighting some aspects of this particular type of modern wireless networks. Related definitions of the vision, the basic element and the architecture of the network are introduced. Moreover, the paradigm applications are introduced and the related wireless access technologies are depicted. Since there is a peculiar characteristics and requirement, we discuss the overall network requirements and challenges that need to be satisfied to reach such an effective network.

The next chapter overviews the related works in the field of IoT communication and topology management. Particularly, we focus on the MAC and clustering techniques, been commonly, used in wireless networks.

2

State of the Art

2.1/ Introduction

The design of an efficient communication system for large scale IoT network is a serious challenge due to the limitation and characteristics of such networks as discussed in the precedent chapter. Therefore, researches have turned to the implementation of new techniques and mechanisms that ease and enable an efficient management of communication channels, stable topology structure and effective data routing in the network. Besides the large number of deployed IoT devices, an extensible and dynamic hierarchic structure can achieve an effective and efficient management [32]. This chapter focuses on two sub-layers directly implied in communication and data aggregation in wireless network, notably, the MAC and Network layers.

The MAC layer needs to be considered to manage the access to the shared communication medium in order to prevent interference and, consequently, improves the transmission reliability and the service QoS. Accordingly, an efficient MAC protocol design needs to deal with the entire feature that impacts the performance, such as the congestion phenomena. For instance, the usage of TDMA-based channel access technique reduces the interference effect. However, slots may be allocated to idle nodes which degrades the system capacity [35].

On the other hand, the network layer is responsible of the structuring of the network topology to ensure a coherent data routing. We focus on clustering techniques because they represent an important and attractive solution used in conventional wireless networks for structuring the network and optimizing its services. The network clustering significantly enhances the performances of numerous applications compared to the conventional flat structure. Moreover, the clustering architecture enhances the interference avoidance due to a better channel allocation. Indeed, a well-structured network allows a better control of communications interference [29].

The present chapter provides a review of the MAC and clustering protocols present in the current literature. First, we introduce a taxonomy of MAC protocols for wireless networks and we highlight the main approaches in this field. Mind that we cannot review the complete research works in detail for this domain. Hence, we rather focus on the most common MAC approaches.Secondly, we propose a classification for clustering algorithms in wireless networks with a detailed description of each algorithm. Thirdly, we highlight the relation between network and MAC layers and we present an overview of cross-layer approaches with a presentation of some relevant approaches in this field. Finally, we discuss the inconvenient of the layered model and the issues of the cross layering.

2.2/ MAC PROTOCOLS CLASSIFICATION

The medium access control (MAC) layer has an important task for ensuring the successful transmission in the wireless network. The main task of the MAC protocol is to manage and coordinate the access to the communication channel while reducing the collisions and interference's among nodes. Communication protocols suitable for MAC protocols can be classified into two main categories, notably, contention free and contention based protocols, as shown in fig 2.1.

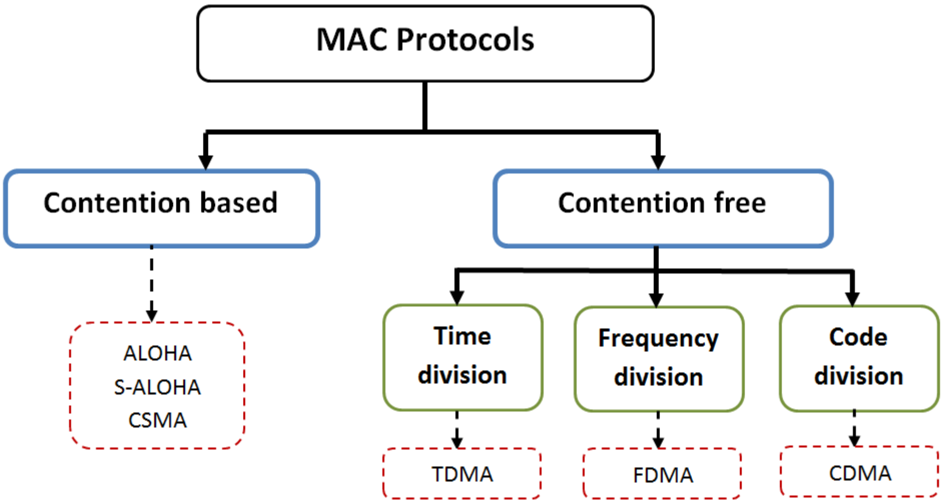


Figure 2.1: MAC protocols classification

2.2.1/ CONTENTION BASED PROTOCOLS (CBP)

The contention based protocols need to accept the channel access before the transmission of each frame. It allows the usage of the radio channel without a pre-coordination. In this type of protocol there are no predetermined time slots to send or receive data. The most used contention mechanism is based on RTS / CTS (Request To Send/Clear To

Send) control packets. Before transmitting, the sender node sends an RTS to its neighbor. If the recipient neighbor is available to receive the data, it return a CTS to the sender. ALOHA, slotted ALOHA [67], CSMA [68] and IEEE 802.11 [45] are some examples of CBP based MAC protocols. The IEEE 802.11 [45] acquired a wide approval due to its simplicity, robustness and flexibility, it is built on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) [68].

ALOHA [67] is a basic communication protocol where nodes transmit their data whenever they have a frame to be transmitted. If the transmission fails, the sender will re-transmit the frame. However, when the communication volume increases, collisions become frequent which reduces the efficiency of the network. CSMA [68] come as an improvement of the basic ALOHA, it prevents collision by using the "listen before talk" operating procedure and senses the shared communication channel before transmitting. However, the transmitter can only sense the shared medium within its transmitting range. Therefore, transmissions out of the range of the sender cannot be detected (hidden node problem).

Contention based protocols are advantageous for sparse networks and support ad hoc communications. However, this kind of protocol is not intended to support multi-hop wireless communication [69]. Indeed, the random access scheme deployed in the standard still suffers from the hidden terminal problem which may cause interference and collision problems.

2.2.2/ Contention free protocols (CFP)

In contrast, Contention-free MAC protocols reserve the channel to ensure a conflict-free packet exchange. Basically, these protocols follow two phases: a random channel access to gain the channel reservation and a scheduled transmission of data packets. With this type of protocol there is no competition to access the communication channel. Each node has a private time/frequency/code that allow it to access the shared medium and no other node in its vicinity will be able to use the same time/frequency/code. A MAC protocol is said contention-free if the transmissions do not collide.

Among the protocols in this category, we distinguish three basic channel access techniques:

- **Time Division Multiple Access (TDMA)** where the time is slotted and each node has a dedicated interval of time during which he can access the channel to transmit its data. Outside this time it is not allowed to access the channel.

  TDMA technique assumes that all the network devices are time-synchronized in some way. However, time synchronization may be infeasible in large scale sensor networks, and it is better not to rely on synchronization in the design of MAC protocols. The TDMA scheme is optimal in terms of throughput for fully connected networks [70]. As we attempt to reduce the overall communication latency. In this work (chapter 3), we focus on the distributed TDMA-based solutions.

- **Frequency Division Multiple Access (FDMA)** allows multiple users to communicate through a single communication channel simultaneously by partitioning the channel bandwidth into several non-overlapping frequency sub-channels and attribute each sub-channel to a distinct user.

- **Code Division Multiple Access (CDMA)** where multiple transmitters can send information simultaneously over a single frequency sub-band. The CDMA optimizes the use of the available frequencies as it uses the entire frequency bandwidth without need for guard sub-bands. CDMA technique uses a spread spectrum technology and a special coding scheme (each transmitter uses an assigned code). The transmitter encodes its data with its own code. On the other hand, the receiver uses the transmitter code to extract its concerned data.

2.3/ Routing protocols

Several classification of the routing protocols exist in the literature, notably, for the flat routing and hierarchical clustering [23; 27; 26; 28; 32; 31; 49; 33; 29; 71].

2.3.1/ Flat routing

A network is said flat [1] when all of its interconnected elements have the same features and the network is not divided into parts or groups. With this type of routing (also known as Data-centric routing), all the network devices have the same functionalities and they collaborate to relay the data. The device decision to route data packets to the destination depend on its location. AODV [72] (Ad hoc On Demand Distance Vector), SPIN (Sensor Protocols for Information via Negotiation) [73] and Directed Diffusion [74] belongs to this class.

The flat routing is convenient for small scale network and offers a high reliability. However, it does not support the dynamicity of the network topology and it does not perform well when the density of the network increases [75].

[1] A flat structure refers to the absence of hierarchy provided by clustering operations.

2.3.2/ Hierarchical routing

In this thesis, we focus on the management of the communication channels in large scale network. Therefore, we concentrate on the network topology management and we addressed the hierarchical clustering to provide an easily manageable network. In the following we present a review of the clustering in wireless networks.

2.3.2.1/ Clustering in IoT: An overview

For large wireless networks, topology management is indispensable to balance the network load, increase the scalability and prolong the network lifetime. As an important network structuring technique employed in the conventional wireless networks, the hierarchical clustering presents an interesting solution to facilitate the organization and considerably improve the performances of numerous applications compared to conventional flat structure [32]. The goal of this technique is to organize the network nodes into small groups according to various parameters and metrics.

2.3.2.2/ Initial clustering works

The Lowest-ID algorithm [76] is the first notable research work on clustering. The main goal is to assign a unique identifier to nodes. The node identifiers are then used to select the cluster heads in such manner that the lowest ID is chosen as a cluster head. The drawback of this algorithm is the recurrent selection of the same nodes as cluster heads leading to unbalanced energy consumption. The Highest-Degree algorithm [77], Least Cluster Change [78] and MOBIC (The first signal strength based clustering algorithm) [79] were proposed later, attempting to introduce new topological and mobility metrics to the clustering process.

The previous described algorithms are based on the 1-hop clustering. A novel algorithm was given in [80] called Max-Min. It is the earliest proposal for multi-hop cluster-based topology. Although these algorithms appear obsolete compared to the recent literature. The majority of current proposals in clustering are based on the set of rules described in these first proposals combined with more novel clustering metrics.

2.3.2.3/ Clustering concept

Although clustering techniques are diverse, grouping nodes into small groups (called clusters) is a shared concept. This structure can fulfill a several network requirements, such as scalability, load balancing, network stability, Quality of Service (QoS), etc. Clustering implies to assign particular roles to nodes, notably, the cluster heads (CH), the cluster members (CM), and the cluster gateway (CG).

- **Cluster head (CH)**: represents the leader of a cluster. It is responsible for relaying information between nodes within the same cluster (intra-cluster communications) or between adjacent clusters (inter-cluster communications). It can communicate with the BS directly or through other CHs. The CH bears some extra tasks compared to other nodes, such as data aggregation and channel access management, which requires more energy consumption.

- **Cluster Members (CMs)**: are ordinary nodes belonging to a given cluster. The corresponding cluster of a node depends on the node's characteristics. These nodes are limited to the tasks defined by the targeted application, for instance sending their sensed data to the CH.

- **Cluster Gateway (CG)**: particular CM node that provides inter-cluster communication with neighboring clusters and transfer data between them. It's usually positioned at the cluster's borders.

The usual method consists of selecting the CH according to a predefined metrics, and then the rest of the nodes join one of the surrounding CHs. CMs that interconnect clusters become gateways, figure 2.2 illustrate the clustering procedure.

2.3.2.4/ Clustering performance metrics

In this section we discuss how clustering efficiency could be evaluated. Indeed, Clustering algorithms are diverse and designed for different scenarios. Therefore, determine which algorithm is the best is not straightforward. However, it is possible to state which algorithm is more suitable for a particular scenario. Accordingly, it is important to determine which metrics need to be evaluated to assess the performance of a clustering algorithm. Fig 2.3 highlights the most common metrics used for performance evaluation of clustering algorithms. These metrics can target general evaluation of the clustering stability, overhead, connectivity and more specific measurements related with the network lifetime, energy efficiency, etc.

2.3.2.5/ Benefits and costs of clustering

This part discusses the relevant benefits of clustering and highlighting the costs of applying such a technique.

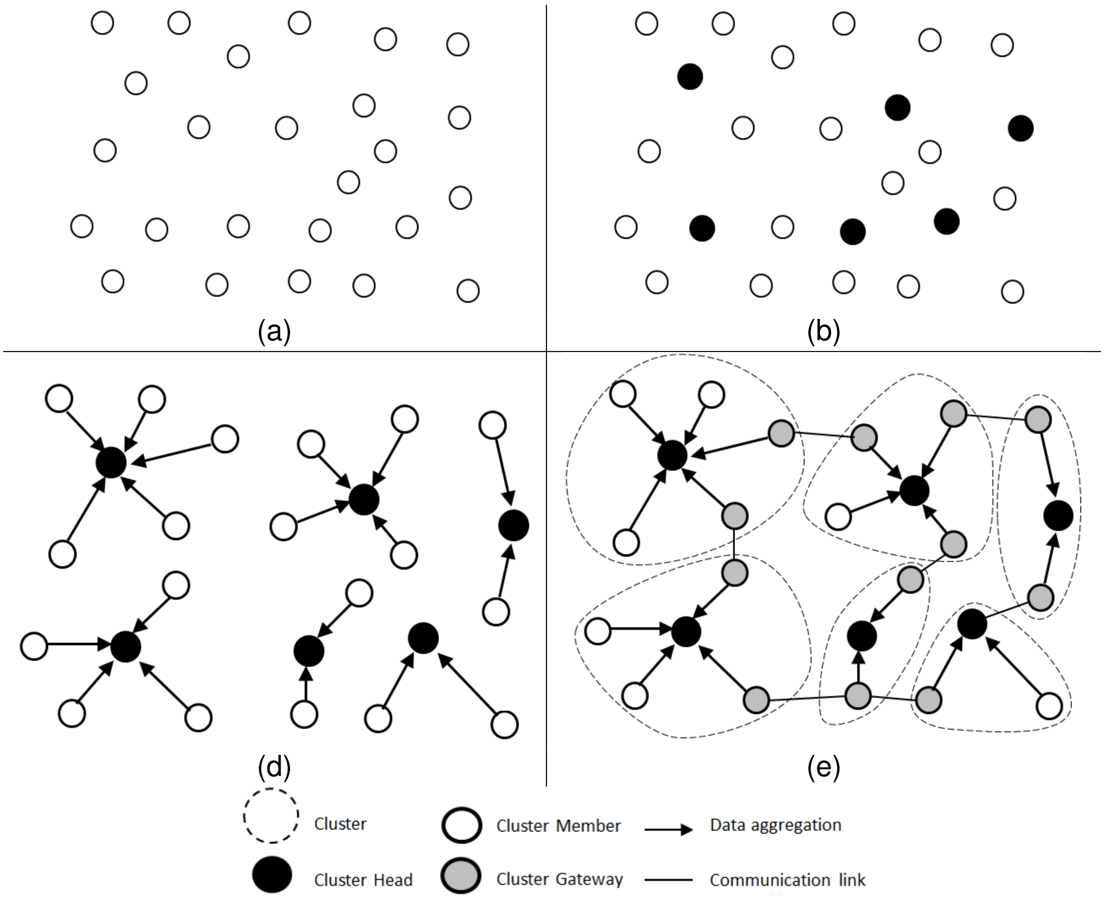


Figure 2.2: Illustration example of the clustering process: Converting a flat network topology (a) to a hierarchical clustered topology (d) and the roles assigned to nodes in the clustered network

Benefits of clustering

In the beginning, clustering technique emerge as a solution for routing scalability problem in ad hoc networking [29]. Indeed, the majority of ad hoc routing proposals were adopted from the IP-based wired protocols, which have shown some difficulties in integrating big flat ad hoc networks. The division of the network into clusters eases the management of the topology and reduces the coordination messages between nodes execution load of the high-level protocols. Additionally, in terms of radio transmission, the prioritization of access to channels it possible to limit access collisions and interference.

The benefits obtained from clustering architectures depend on the application. For example, fast channel allocation and better intra-cluster coordination represent the main advantages for MANETs, whereas the Scalability of routing protocols and the energy efficiency are considered as the main benefit of WSNs. In the new proposed scenarios, the objective is to favor data aggregation by using stable links between cooperating nodes

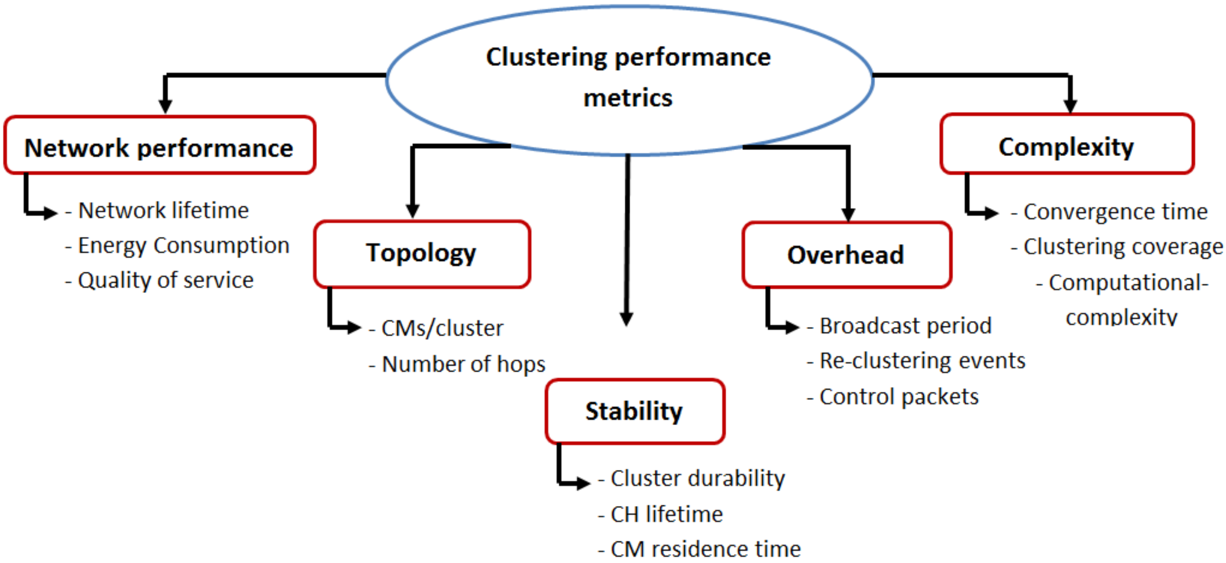


Figure 2.3: Common metrics adopted in performance evaluation of clustering algorithms

and better coordination to provide contention free intra-cluster communications. Moreover, the clustering architectures can considerably reduce routing complexity by reducing the number of routes within the network, which makes the network more scalable. Another advantage of the clustering is the interference and collisions avoidance provided by a better channel allocation. Indeed, a good organization of clusters allows an easier deployment of non-interfering channels and favors the network load balancing.

Costs of clustering

As all the other techniques, Clustering also has its own costs including the messages overhead, the energy consumption and the algorithm complexity. For instance, explicit control messages need to be exchanged to perform the hierarchical topology, these messages may include the topological information (e.g. list of neighbors, quality of the links, etc.) or node related information (e.g. mobility, battery levels, hardware capabilities, etc.). These packets are used for neighbor discovery, cluster head selection, cluster formation and maintenance and also for routing tasks. The exchange of these messages includes an additional overhead and energy consumption. Moreover, the complexity of the clustering algorithm must also be taken into account. Indeed, the usage of this technique engenders additional computational resource requirements and increase the time required by the algorithm to complete the hierarchical topology. When the clustering process takes a considerable time, the information exchange can get outdated and obstruct the network performance.

2.3.3/ Clustering taxonomy

According to the current literature [23; 27; 26; 32; 31; 29], there is a number of categories according to numerous metrics and criteria used for the CHs election. In this part we follow the classification provided in [29]. We include, however, more novel approaches. Therefore, we propose to classify clustering algorithms into the following categories: Mobility-aware clustering, Energy-Aware clustering, Load-balancing clustering, neighbors based clustering and multi-criteria based clustering (as illustrated in figure 2.4). We detail the main innovations in these categories and address some incremental research efforts that improved these approaches. In the following, we discuss the different clustering algorithms based on the proposed taxonomy.

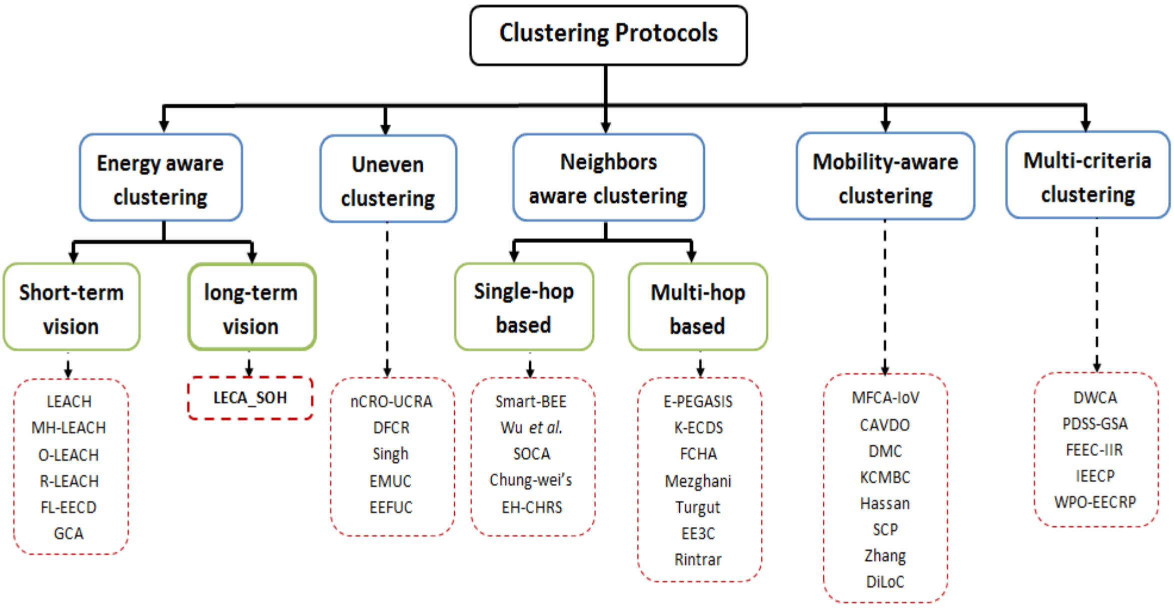


Figure 2.4: Taxonomy of clustering algorithms in wireless networks.

2.3.3.1/ Mobility based clustering

Mobility-based clustering approaches treat the stability of the clustered topology in mobile environments. Improving the stability refers to maximizing the connectivity among cluster members, i.e. enhancing the cluster lifetime and residence time of nodes within the clusters. In this type of clustering, mobility information (relative velocity, acceleration and direction) are used to anticipate nodes movement and form the clustered topology. Stable clusters are formed by nodes that move with similar features (direction and speed). These approaches can be classified according to the used source information. The major approaches are mostly presents in VANETs networks. Nodes are supposed to be equipped with a GPS or a distributed localization system to gather mobility data about speed, direction, position, travel route, among others. Mobility clustering includes also the environmental-based approaches that use diverse information provided by the neighborhood, such as received signal-strength (RSS) to predict the connection time between nodes. The environmental-based approaches are designed for networks where nodes are deprived of the GPS, i.e. wireless sensors networks.

Authors in [81] proposed a new mechanism based on Moth Flame Clustering optimization (MFO) for Internet of Vehicles, called MFCA-IoV. The protocol optimizes transmission robustness. In [82], Aadil et al. proposed a new clustering algorithm for Internet of vehicles based on Dragonfly Optimizer (CAVDO). CAVDO uses a Mobility Aware model (MA-DTR) to calculate a dynamical transmission range. The proposal improves the topology stability in dynamic environment by using a swarm-based multi-objective method to structure the network.

Authors in [83] presented an efficient Dynamic Mobility based Clustering (DMC) algorithm to form a stable core network for data aggregation and propagation. The clustering process is performed by taking into consideration the moving direction of devices, the relative speed, average distance and links stability. The approach uses a temporary cluster head technique to enhance the cluster formation phase and uses a safe distance threshold to monitor clusters size. A similar approach is introduced in SCP (Stable Clustering Protocol) [84], where the clustering algorithm uses a metric based on the nodes' speed. Network nodes estimate their speed vector using GPS and compute the average speed of the neighborhood using the received information. The stability rate is computed by comparing the node speed with the average speed of it surrounding neighbors. In order to select stable nodes as CHs, a willingness parameter is calculated based on node's average velocity and the number of neighbors with similar velocity. The node with higher willingness parameter is selected as a CH. The algorithm uses a threshold value that serves to evaluate if a CH needs to resign its role. Therefore, the CH that leaves the cluster can resign before losing the connection with the CMs, which accelerates the re-clustering process.

Hassanabadi *et al.* in [85] presented a mobility aware clustering algorithm for VANETs that adopts the concepts of data mining by classifying data samples into groups of similarity. The approach uses an affinity propagation technique to form mobile clusters where nodes exchange messages to follow the affinity propagation. The clustering process uses a similarity metric that takes into account the node position and mobility. The approach is fully distributed and ensures clustering convergence within a predefined time. KCMBC (K-hop Compound Metric Based Clustering) approach [29] diverts from the idea of using mobility and location information, the clustering process uses a prediction measure to estimate the link expiration time. The link expiration time for a node is measured according to the speeds of neighboring nodes, the node that maintain more stable links is elected as cluster head. The cluster topology is formed following the same approach used in [80] for the k-hop clustering. KCMBC uses a dynamically adaptive broadcast period to transmit the data to nodes according to their mobility. KCMBC adapts the broadcast period according to the variability of the network.

Zhang et al. proposed in [86] another relative mobility-based clustering algorithm for VANETs. The proposal uses the ratio of transmission delay between two consecutive packets to calculate the devices k-hop relative mobility. The aggregate relative mobility value is the summary of the relative mobility times value of all neighbor nodes within k-hop. The device that has the smallest aggregate relative mobility is selected as the CH. The authors in [87] proposed a Distributed and Location-aware Clustering approach called DiLoC, where the CHs (anchor nodes) are established before the clustering process. These anchor nodes are responsible for initialization of the clustering process. DiLoC is based on proximity election, relying on devices scattered along the network to determine location information for the entire cluster structure. i.e. unassigned nodes (nodes that have not yet received their assignment to CM or CH) that receive a join message from an anchor node (regardless of the number of hops) can join the cluster. As the CHs are selected before the clustering process, the clusters cardinality is limited, which may restrict the scalability in case of a dense networks. Table 2.1 outlines a brief comparison between the mobility based clustering approaches presented in this section.

Table 2.1: Comparison of mobility based clustering algorithms.

Algorithm	**Topology**	**Clustering structure**	**Energy consid.**	**Scalability**	**Lifetime improv.**	**Clustering parameter**
MFCA-IoV [81]	Central	Single-hop	Low	Medium	Low	Moth flame opt.
CAVDO [82]	Central	Single-hop	Low	High	Low	Dragonfly opt., mobility
DMC [83]	Central	Single-hop	Low	Medium	Low	Direction, Position, Links lifetime
KCMBC [29]	Distributed	Multi-hop	Low	Medium	Low	Mobility
Hassan. [85]	Distributed	Single-hop	Low	Low	Low	Position, Mobility
SCP [84]	Central	Single-hop	Low	Low	Low	Speed
Zhang [86]	Central.	Multi-hop	Low	Hight	Low	Mobility, Connectivity
DiLoC [87]	Central.	Multi-hop	Low	Low	Low	Location

2.3.3.2/ Energy and load balancing aware clustering

Energy-aware clustering is mostly applied with resource restrained wireless networks, such as wireless sensor networks, where the mobility is not the main focus. The primary concern in these networks is exclusively devoted to increase the network lifetime which is typically defined as the time elapsed before the first network node drains out of its battery (short-term vision). A considerable amount of literature addresses the lifetime improvement of wireless networks [28; 31; 33; 32; 26; 2; 8].

Low Energy Adaptive Clustering Hierarchy (LEACH) [88] is one of the most pioneering and the well-known algorithms in this field. It represents a distributed cluster-based approach that use randomized rotation of CHs relied on a probabilistic threshold to distribute the energy load among network nodes. LEACH renders a considerable energy saving and prolongs the network lifetime compared to the flat clustering approaches. However, the primary drawback of this approach is the possibility of selecting CHs with low residual energy, which may results in a quick death and deterioration of the network performance. LEACH attracted the attention of the research community working in the area of wireless network. Several approaches come up with various improvements of this protocol to alleviate the energy problem (e.g. MH-LEACH [89],O-LEACH [90], R-LEACH [91] and others [92]).

A Multi-Hop version of LEACH protocol (MH-LEACH) was introduced in [89]. MH-LEACH proposes a multi-hop architecture within each cluster to promote energy preservation. Some specific CM nodes are considered as intermediate nodes to relay the collected data toward the CH. In [90], the authors combine LEACH protocol with Genetic Algorithm (GA). LEACH component determines the clusters, while the GA component is used to locate the optimal routing paths using a fitness function. This latter uses the residual energy of the nodes for the selection of the CHs. O-LEACH outperforms the classic LEACH by improving the packet delivery rate, throughput, and energy saving. Residual energy-based LEACH (R-LEACH) protocol [91] selects the CH by combining multiple parameters, such as the remaining energy and the optimum number of network CHs. The approach applies the same mechanism for the CH election as LEACH. However, the new CH is elected at the end of the round based on the residual energy. In [93], Hamzah et al. proposed a new clustering method called FL-EECD based on the minimum separation distance between CHs. A fuzzy inference system is used to acquire better parameters. The CHs are selected based on the residual energy, location suitability, density, and distance from the BS. The approach sets a minimum separating distance between CHs and uses a Gini index [93] to estimate the energy efficiency of the clustering approach to ensure an even distribution of energy through nodes. This model aims to select an optimal set of CHs. However, it does not consider the load balancing of the aggregated data among CHs.

The authors in [94] proposed a Grid-based Clustering Algorithm (GCA) to enhance load balancing and extend the network durability. The GCA analyzes the energy consumption during the network execution and based on this parameter, data packets distribution takes place. Polynomials metrics valuate the optimal cluster size to reasonably balance the energy consumption. In this approach, the network is divided into several grid clusters based on the optimal cluster size, the network load distribution is specified accordingly.

Table 2.2 illustrates a brief comparison of the energy aware clustering approaches presented in this section.

Although a considerable amount of research works focus on the lifetime extension of wireless networks using energy aware approaches. The proposed clustering protocols, up to this point, have been mostly based on devices residual energy or the State of Charge (SoC) optimization for non-rechargeable batteries. These solutions focus on maximizing the duration of batteries discharging (short-term vision) and ignore the impact of operational activities of the devices on the battery State of Health (SOH). Therefore, such approaches are not adapted to rechargeable batteries optimization (long-term vision).

Table 2.2: Comparison of energy aware based clustering algorithms.

Algorithm	**Topology**	**Clustering structure**	**Energy consid.**	**Scalability**	**Lifetime improv.**	**Clustering parameter**
LEACH [88]	Distributed	Single-hop	Low	Low	Low	Load balance
MH-LEACH [89]	Distributed	Multi-hop	Moderate	High	Moderate	SoC, reliability
O-LEACH [90]	Distributed	Multi-hop	High	Moderate	High	SoC, latency
R-LEACH [91]	Distributed	Single-hop	high	Moderate	Moderate	SoC, density
FL-EECD [93]	Central.	Single-hop	Moderate	High	Moderate	SoC, position, density, Distance to BS
GCA [94]	Central.	Single-hop	Moderate	Low	Low	SoC

2.3.3.3/ Energy aware protocols for rechargeable batteries

The IoT applications are constrained by the required devices' energy consumption. Batteries, commonly, represents the only source of energy of the IoT devices and,therefore, defines the nodes lifetime. This problem undermines the integration of the traditional wireless protocols in the IoT networks[95]. The rechargeable battery concept has attracted much attention (e.g. Energy Harvesting WSN and Rechargeable WSNs, EH-WSNs or R-WSNs) with the promise of a sustainable IoT networks [95; 96]. This kind of batteries benefit from ability to be recharged many times [97] using techniques such as piezoelectric generators, RF harvesters, solar panels and wind turbines. The idea is to convert energy sources, such as body heat, foot strike, finger strokes, wind and sunbeams into electricity [95]. However, the durability of rechargeable batteries is not unlimited and omitting the degradation aspect may result in a high replacement and maintenance costs. Therefore, saving battery state of health (SoH) is mandatory even with rechargeable batteries (lifetime improvement in the long-term).

Energy-aware clustering protocols typically assume that the battery is non-rechargeable or the degradation impact on the battery SoH is neglected. Whereas, rechargeable batteries are nowadays a commonly used technology and the contactless harvesting solutions become reality [98]. In this part (section 2.3.3.3), we propose an original approach that integrates rechargeable battery degradation within the networking protocols.

Although, the typical part of research work in wireless networks focus on energy limited devices, some current proposal addressed the problem of energy optimization for rechargeable nodes in wireless networks [96]. For instance, Lin et al. [99] considered the problem of maximizing the data collection rate in rechargeable wireless sensor networks (R-WSN). In this proposal, the problem of data aggregation in multisource/multi-sink sensor networks is formulated as a linear programming problem where the devices have to consume less energy than they collect. Authors in [100] deal with the scheduling of the recharging phases by considering the temporal and space constraints in R-WSN. In this approach, a mobile charger moves to fulfill the charging requests. Arivudainambi et al. [101] considered the problem of wireless chargers location in rechargeable wireless sensor networks. The work in [102] presents an adaptive MAC protocol for R-WSN (Rechargeable WSN). The goal is to specify a method that enables each node to measure the optimal maximum and minimum charging thresholds according to the number of exchanged data packets. When the remaining energy of a given node reaches a minimum threshold, the node trigger the recharging process until the maximum energy threshold is reached. During the recharging process, the node cannot send or receive data.

2.3.3.4/ Uneven clustering

In multi-hop clustering, nodes closer to the Base Station (BS) exhaust their energy rapidly compared to other nodes. This is due to the transmission of their own data in addition to the aggregation of other nodes packets. This causes irregular energy consumption over the network. Therefore, nodes closer to the BS tend to die quickly leading to a network disconnection. This type of problem is known as hot spot problem. Unequal clustering is one of the proposed solutions to overcome this issue [103]. It brings additional advantages like extending the scalability and the durability of the network [104]. Several uneven clustering protocols have been proposed in the literature. [105; 106; 103; 107; 104]

In [108], the authors presented a distributed approach for unequal clustering inspired from a chemical reaction optimization paradigm called nCRO-UCRA. The approach aims to eliminate the hot spot problem by merging the unequal clustering with a Linear Programming formulation for the CHs selection. The network is partitioned into unequal clusters where small size clusters are located near to the BS. The CHs forward their data to the BS by using the UCRA algorithm. The approach aims to achieve better performance in terms of residual energy and network lifetime.

A distributed fuzzy logic-based unequal clustering approach (DFCR) is proposed in [109] to solve the early death problem of nodes close to the BS. DFCR forms the clusters of unequal sizes where the size of each cluster is calculated by a distributed fuzzy logic approach. The CHs are selected according to their distance toward the BS and their residual energy. After the clustering, a virtual backbone of CHs is constructed and network nodes use a cost function to determine the shortest backbone path to route their data.

Singh et al. [110] presented a multi-hop intra-cluster technique to attenuate the risk of hot spot and avoid intra-cluster communication problems. This scheme uses an unequal grid based clustering to divide the network into unequal and fixed square shaped clusters. Clusters size is fixed during the entire network lifetime to avoid going through clusters formation in each round and save energy. The proposal assumes both centralized and distributed methods to structure the network. Clusters formation and CHs selection are achieved centrally, whereas intra and inter cluster communications are performed in a distributed fashion. The selected CHs send their data to the BS by using multi-hop inter-cluster links. If the remaining energy of a CH drops under a specific fraction compared to it neighbors, the BS triggers a new clustering process. The aim of the approach is to reduce the communications distance and reduce the CHs rotation at each round.

An Energy-efficient Multi-hop routing with Unequal Clustering approach (EMUC) is presented in [103]. EMUC aims to distribute the energy dissipation within the network. This scheme extends the life of the network by reducing the energy cost of long distance transmission and approving multi-hop communication to aggregate the collected data. It uses the unequal clustering to reduce the effect of the hot spot problem. The CHs election is distributed and multiple tentative CHs are selected according to a probabilistic model and compete to be CHs based on their residual energy. Inter-cluster communications are conducted in multi-hop to reduce the energy consumption of CHs. Each CH selects the appropriate routing path to relay its data by taking into consideration a compromise between the energy of relay nodes and the energy cost of relay links. The work in [107] illustrated a new distributed communication architecture using multi-hop transmissions and a fuzzy logic mechanism to provide energy-efficient communication with heterogeneous clustering (EEFUC). The proposal is based on four stages, namely, competition radius determination, CHs election, cluster joining, and the selection of relaying paths. The fuzzy logic is applied during the CHs election by considering various input parameters such as residual energy, distance to the BS and the network density. The authors investigate the application of the fuzzy logic using various parameters and functions to obtain an appropriate clustering criterion.

Table 2.3 outlines a brief comparison of the uneven clustering approaches presented in this section.

Table 2.3: Comparison of uneven clustering aware algorithms.

Algorithm	**Topology**	**Clustering structure**	**Energy consid.**	**Scalability**	**Lifetime improv.**	**Clustering parameter**
nCRO-UCRA [106]	Distributed	Multi-hop	Moderate	High	Moderate	SoC, position, density, distance to BS
DFCR [109]	Distributed	Fuzzy-based	Moderate	Moderate	Moderate	SoC, position, distance to BS
Singh [110]	Distributed	Multi-hop	High	Moderate	Low	SoC, load balance
EMUC [103]	Distributed	Multi-hop	Moderate	High	High	SoC, reliability
EEFUC [107]	Distributed	Multi-hop	Low	Low	Moderate	SoC, density, distance to BS

2.3.3.5/ Neighbors aware clustering

Clustering approaches in dense and low mobility networks usually focus on balancing clusters size alternatively with the stability of the clustering topology. The Cluster size represents a trade-off between the inter-cluster and intra-cluster communication performance [29]. Indeed, big clusters have a significant effect on the CHs energy consumption, therefore, different clustering proposals try to enhance the network performance by forming clusters with optimum size. Proposed clustering approaches can be classified into two main categories according to the cluster radius. Namely, the single hop (or 1-hop) based clustering approaches where each node is one hop away from its relative CH (directly connected). The second category regroup the multi-hops (or k-hop) base clustering approaches where network nodes may not be directly connected with their CH.

1-hop based clustering

Authors in [40] proposed a smart and balanced one-hop clustering algorithm (Smart-BEE) for IoT Networks. The authors approached the clustering scheme from another perspective, they accommodate the multiple communication interfaces enabled in 5G networks using Multiple-In and Multiple-Out (MIMO) communications to perform energy efficient topology, maintain the network coverage and improve the Quality of user Experience (QoE). The proposal is centralized which reduces the performance in case of a high dynamic network and makes it sensitive to the central point of failure issue. Wu *et al.* [111] use an ant colony optimization to provide an energy efficient clustering approach. The proposal aims to find a minimal set of CHs and preserve the battery SoC by load-balancing the number of devices among network clusters.

The authors in [112] proposed a new self-organized Clustering Algorithm (SOCA) for VANETs. The approach focuses on the single-hop clustering to reduce the network overhead while maintaining the communications security using asymmetric cryptography. The

CHs election uses an improved algorithm of Independent Set Problem with a secret-key agreement technique. Authors in [113] proposed an energy harvesting CH rotation scheme (EH-CHRS) for green wireless networks based on an analytical double chain Markov model. The main motivation of the approach is to perform a CH rotation and minimize the energy overflow during high data traffic by assigning the CH role to nodes that consume less energy during the data aggregation phase.

Chung-Wei's *et al.* [111] proposed a multi-objective ant-colony optimization algorithm for cluster formation. The proposal considers three objectives, notably, minimize the number of cluster heads, balance the number of members among cluster, reduce the distance between nodes of the same cluster. The algorithm is centralized and aims to find the Pareto optimal set of cluster heads [111] and associated CMs that minimize the energy dissipation. Table 2.4 shows a brief comparison of the single-hop clustering approaches presented in this section.

Table 2.4: Comparison of single-hop clustering aware algorithms.

Algorithm	**Topology**	**Intra-cluster structure**	**Energy consid.**	**Scalability**	**Lifetime improv.**	**Clustering parameter**
Smart-BEE [40]	Central.	Single-hop	Moderate	Low	Moderate	SoC, load balance
Wu *et al.* [111]	Central.	Single-hop	Moderate	Low	High	SoC
SOCA [112]	Central.	Single-hop	Low	Moderate	Low	Cryptography
EH-CHRS [113]	Central.	Single-hop	Moderate	Low	Moderate	SoC, load balance
Chung-Wei's [111]	Central.	single-hop	Moderate	Low	Low	Ant-colony opt., load balance

K-HOP BASED CLUSTERING

The single hop intra-clustering offers a suitable solution to structure the network. However, for extensive networks with short range radio signals, a significant number of clusters are formed. Accordingly, isolated devices may arise and connectivity may be difficult to guarantee. Hence, Multi-hop (or k-hop) intra-clustering produces a consistent structure in a dynamic large scale networks. Indeed, increasing the number of hops inside network clusters implicitly decreases the number of reconfiguration events that may occur (CH election and nodes re-affiliation) [29], which improve the network stability.

E-PEGASIS [114] is a multi-hop intra-clustering mechanism that aims to reduce the redundancy of data transmitted toward the BS. E-PEGASIS compute a dominating set in the network and employs an ant colony optimization to select the sub-optimal routing chain from the dominating sets. This approach provides an energy efficient solution for routing the network data. However, it requires an extended time complexity to organize the nodes into chain. Authors in [115] proposed a k-hop Energy Constrained intra-clustering technique based on the Dominating Sets theory called K-ECDS. The proposed distributed approach considers the energy limitation and models the cluster heads election problem based on the quality of communication and nodes cardinality.

Authors in [106] presented a centralized energy efficient fuzzy-based cluster head amendment (FCHA) approach to increase the lifetime of the network. Clusters topology is performed dynamically by the BS. The CH role rotates periodically and the clustering applies an adjacency matrix and a Fuzzy Logic (FL) approach for the CHs election. The FL concept specify nodes centrality and residual energy as an input parameter. FCHA uses an actuation point as a factor to decide the periodicity of the current cluster head.

Mezghani [116] proposed a distributed multi-hop intra-clustering approach where the network is divided into k-hop large clusters. The CHs election is distributed and weight based on the residual energy, nodes degree and the communication quality. The cluster topology is achieved using the triangulation theory of Khalimsky to ensure an effective intra-clustering routing and reduce the energy wastage.

Authors in [117] presented a multi-hop intra-cluster clustering architecture. Initially, nodes enter a sleeping period and compute a wake up time slot according to their degree and the average distance to their neighbors. Nodes with a large set of neighbors and low average distance are awakened in an earlier execution time to be privileged for the CH election. Although the CHs selection provides more energy efficiency and balanced cluster formation, the clustering time is prolonged.

Authors in [118] proposed an Energy Efficient intra-clustering technique (EE3C) where multiple high energy nodes act as cluster heads within each cluster instead of a single cluster head. The BS collects node's information to determine their location and lifetime and divide the network into rectangular sectors (clusters) to load balance the energy in the monitored area efficiently. The CH election is centralized at the BS and only one cluster head acts as master CH for a given cluster to gather and send the sensed information to the BS. The master role rotates periodically among cluster CHs after a particular number of rounds. EE3C improves the intra-clustering efficiency, however, the clustering process is fully centralized. Authors in [119] present a reliable energy efficient manner for intra-clustering in wireless networks (RINtraR). The proposal focuses on optimizing the QoS of wireless communication to reduce energy waste. Each node selects the path that consumes the least energy leading to the BS by assigning a transmission quality value to each communication link that relays it with the neighbors. The clustering process of RINtraR is inspired from LEACH which leads nodes with high connectivity to drain faster since they will be used to relay a high rate of data. Table 2.5 exhibits a brief comparison of the multi-hop clustering approaches presented in this section.

Table 2.5: Comparison of multi-hop clustering aware algorithms.

Algorithm	Topology	Intra-cluster structure	Energy consid.	Scalability	Lifetime improv.	Clustering parameter
E-PEGASIS [114]	Central.	Multi-hop	Low	High	Low	Dominating sets, ant colony opt.
K-ECDS [115]	Distributed	Multi-hop	Moderate	High	Moderate	SoC, reliability, cardinality
FCHA [106]	Central.	Multi-hop	Moderate	High	Moderate	SoC, position, load balance
Mezghani [116]	Distributed	Multi-hop	Moderate	High	High	SoC, cardinality, reliability, Khalimsky theory
Turgut [117]	Distributed	Multi-hop	Moderate	Low	Moderate	Cardinality, reliability, TDMA
EE3C [118]	Central.	Multi-hop	Moderate	Low	High	SoC, position
Rintrar [119]	Central.	Multi-hop	Moderate	Moderate	Moderate	SoC, reliability

2.3.3.6/ MULTI-CRITERIA BASED CLUSTERING

Instead of using a single metric for the CHs election, Multi-criteria clustering approaches consider different parameters simultaneously for the clusters formation, such as mobility, location, nodes degree and the energy level. This class of approaches elects cluster-heads that are more convenient to a particular environment in order to meet the requirements of the network. It uses a function that combines the multiple metrics and represents the weight of a node. The weight formula W(i) for a given node i with the metrics $A(i)$, $B(i)$ and $C(i)$ is diffined as in 2.1.

$$W_i = \alpha \times A(i) + \beta \times B(i) + ... + \gamma \times C(i) \ \wedge \ \alpha + \beta + ... + \gamma = 1 \qquad (2.1)$$

where $\alpha, \beta, ..., \gamma$ represent the weighted coefficients of the corresponding metrics. The value of the weighting coefficient presents a main challenge in this category. Indeed, increasing the value of a parameter comes on the expense of reducing the importance of the other parameters. Therefore, the selection of the optimal weighting coefficient vector is hence not a trivial task. These coefficients are adjusted based on the system requirements and the network environment. A particular weight coefficient may be adjusted relatively to the others to reach an optimal result for a particular network configuration. For example, in a low-density environment, the residual energy should be favored. In case of a dense network, the connectivity should be favored. Whereas, in a noisy environment, the communication link quality needs to be considered.

In [120], Essa et al. presented a dynamic weight clustering-based algorithm (DWCA) to equally distribute the network load, expand the network lifetime and strengthen the network scalability. DWCA employs a weight based technique for the CHs election based on the residual energy, the current node degree and the distance toward the BS.

In [121], the authors introduced a fuzzy based energy efficient clustering protocol (FEEC-IIR) for WSN and IoT systems. In this proposal, a multi-criteria decision-making technique is employed and the CHs election is performed based on energy state of charge status (SoC), QoS's impact, and node position to save energy and minimize the packets traffic. Improved Energy-Efficient Clustering Protocol (IEECP) [19] determines the optimal number of clusters for the overlapping balanced clusters. This latter are formed based on a modified fuzzy C-means algorithm, which balances the energy consumption of nodes. The approach uses a CH rotation mechanism combined with a CH back-off timer mechanism to select CHs at optimum locations. WPO-EECRP protocol [122] considers multiple clustering factors for CHs election, namely, the residual energy, the average distance toward the BS, and the number of nodes in the neighborhood. It aims to ensure the scalability of the protocol and to provide a proper clustering control by altering the clustering parameters. The approach is completely distributed and requires the exchange of control packets to estimate the distance separating the nodes and their CH.

In [123], the authors proposed to combine a new version of a gravitational search algorithm (GSA) with a Power Distance Sums Scaling (PDSS) function to determine the optimal number of clusters in each round. This method dynamically uses the position and the remaining energy of nodes to reduce the number of active CHs and decrease the energy consumption. It uses a fuzzy logic controller which considers the residual energy and the transmission link quality to improve the performance of the method. Table 2.6 depicts a comparison between the multi-criteria based clustering approaches presented in this section.

Table 2.6: Comparison of multi-criteria based clustering aware algorithms.

Algorithm	**Topology**	**Clustering structure**	**Energy consid.**	**Scalability**	**Lifetime improv.**	**Clustering parameter**
DWCA [120]	Distributed	single-hop	Moderate	Low	Moderate	SoC, degree, distance to BS, density
PDSS-GSA [123]	Central.	single-hop	Moderate	High	Moderate	SoC, position, reliability
FEEC-IIR [121]	Central.	single-hop	Moderate	Low	Low	SoC, links QoS, position
IEECP [19]	Distributed	multi-hop	High	Moderate	High	SoC, position, loaad balance, Backoff timer
WPO-EECRP [122]	Distributed	multi-hop	Moderate	High	Moderate	SoC, Distance to BS, density

2.4/ Cross-layer aware MAC and routing protocols

2.4.1/ Cross layering definition

Typically, protocol architectures follow particular layering principles to enable interoperability, fast deployment, and effective implementations. However, the poor coordination between layers restricts the performance of the model architectures due to the explicit challenges caused by wireless transmission nature. The layered model inspired from the OSI model, used to ensure network communications, works on the principle of separation layers. Although this principle is effective for wired networks, it is not an optimal solution for IoT networks where other features must be taken into account, such as limited resources and dynamic topology. To address such limitations, the cross-layer design has been proposed. The cross-layer design allows protocols belonging to different layers to cooperate and share their information with other layers in order to efficiently use the network resources and overcome the limits of each protocol. The ITU [2] (International Telecommunication Union) considers that cross-layering consists of allowing interactions between different layers (including non-adjacent layers) in order to achieve gains in network performance.

As the objective of this theses focus on the management of the communication channels, we are interested in cross-layer approaches concerning the Network layer and the MAC layer because their decisions have a significant impact on communication performances.

2.4.2/ Cross layering concept

The main idea is to maintain the functionalities of the initial layers while allowing the coordination, interaction and joint optimization of protocols crossing different layers. Cross-layering can be achieved between two or more layers and it can be based on any combination of layers [124] as shown in figure 2.5. However, the protocols that use the energy information are not considered as cross-layer protocol, although the energy information belong to the Physical layer, it is not considered as a part of a protocol of this layer.

2.4.3/ Classification of cross layer protocols

There is a considerable number of cross layer protocols proposed in the current literature. These protocols can be classified according to the following categories (figure 2.6):

[2] The International Telecommunication Union (ITU) is a specialized agency of the United Nations responsible for all matters related to information and communication technologies. https://www.itu.int/fr/Pages/default.aspx

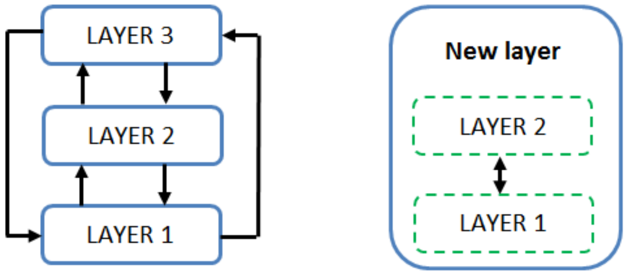


Figure 2.5: Cross layering design

- MAC protocols using information from the routing protocol (Mac-aware routing approaches).
- Routing protocols that consider the information provided from the MAC layer (Routing aware Mac protocols)

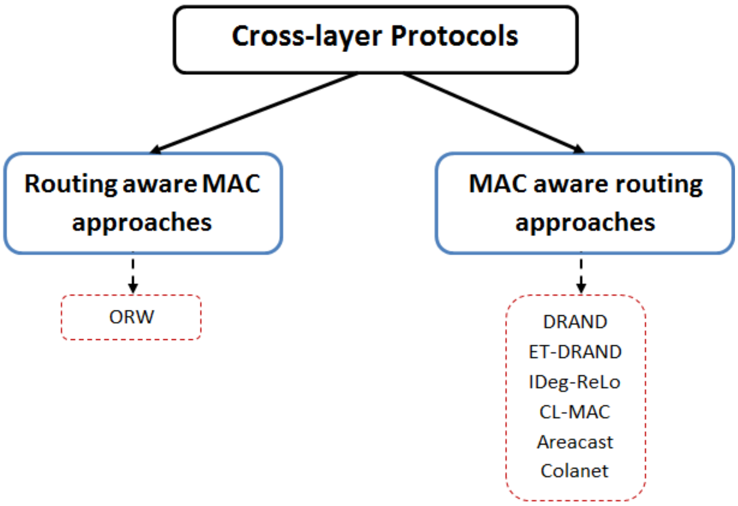
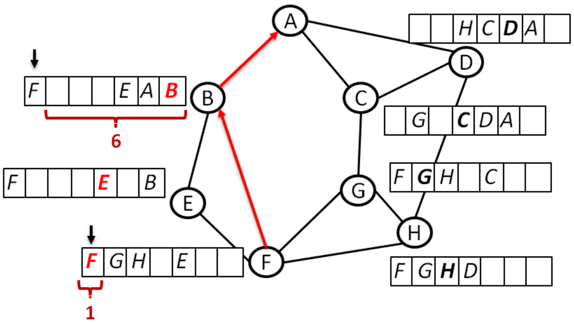


Figure 2.6: Cross layer (MAC-Routing) protocols classification

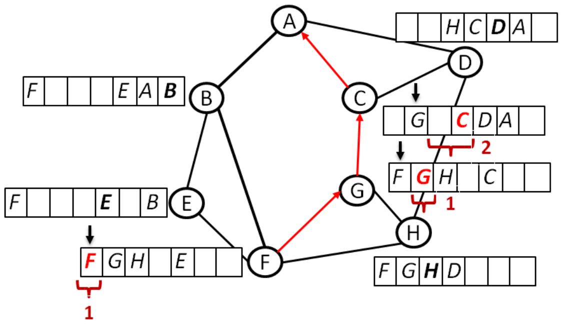
2.4.4/ MAC-aware Routing approaches

Routing in wireless networks has traditionally been divided into two steps: The routing protocol first chooses a next hop, and the MAC protocol then waits for the intended destination to wake up and receive the data. This design makes it difficult to adjust to connection dynamics and causes delays while waiting for the next hop node to wake up. Therefore, the usage of a cross layer approach allows a coordination between these two layers (MAC and Routing), in order to improve the aggregation delay.

To illustrate this category, Fig. 2.7 shows an example of an 8 network nodes (A to H) using TDMA where node F want to aggregate date to node A. With a conventional routing approach without cross-layering (e.g. without considering the MAC-TDMA slots), the routing scheme will route the data toward the shortest path (e.g. toward the node B as in Fig. 2.7(a): $F \xrightarrow{\text{1 slot}} B \xrightarrow{\text{6 slots}} A$). However, with this scheme the aggregation of the data packet will take 7 time slots to reach node A. Whereas, when using a cross layer scheme that considers the TDMA scheduling (e.g. when using a MAC-aware routing approach) the data will be aggregated toward the nodes G and C as shown in Fig. 2.7(b) ($F \xrightarrow{\text{1 slot}} G \xrightarrow{\text{1 slot}} C \xrightarrow{\text{2 slots}} A$). Although, this later scheme involve more hops to reach the node A, this scheme will takes only 4 time slots to route the data to A. Therefore, the usage of this later scheme is advantageous for critical systems as it improves the communication latency.



(a) 2 hops / 7 time slots



(b) 3 hops / 5 time slots

Figure 2.7: Comparaison between (a) conventional routing and (b) MAC-aware Routing approach

In this category, the MAC protocol, once determined, sends it time scheduling information to the routing protocol so that this latter can decide the next hop for each node. This subclass is not widely exploited in the literature. ORW (opportunistic routing for WSNs) [125] belong to this category. ORW is a routing protocol based on MAC protocol information (e.g. the TDMA time schedule) and which aims to decrease delay and power consumption by using all neighbors as potential relay nodes. This approach uses an opportunistic routing based on the concept of duty cycling for each node where the data are sent to the first awaken neighbor. Although ORW is distributed, it requires an iterative process to stabilize the calculation and this procedure becomes more and more difficult with dense networks.

2.4.5/ Routing-aware MAC approaches

Rhee et al. [126] proposed one of the classic algorithms in this category called Distributed and Randomized TDMA Scheduling for Ad Hoc networks (Drand). In this approach, nodes assign their own slot by coordinating requests with their neighbors. Li et al. [70] propose an amelioration of Drand's work, called ET-Drand, by adding the Energy-Topology (ET) factor. They consider the residual energy of nodes and the surrounding neighbor's topology to enhance the distributed scheduling process and improve the energy preservation.

Louail et al. proposed in [127] a centralized cross layer slots allocation protocol based on the routing information, called IDeg-Relo (Interference Degree Remaining Leaves Ordering). This approach allocates a slot to a node based on its position in the routing tree. The algorithm allocates to each node a time slot different from those assigned to interfering nodes. This approach reduces the latency, but faces the central point of failure problem.

Authors in [128] presented a cross-layer scheme designed to improve the reliability of wireless networks under the pretext of inaccurate devices and links called AreaCast. This approach represents a MAC layer protocol that considers the routing information to dynamically avoid faulty nodes and unstable links. AreaCast intends to improve the robustness of routing protocols and the packet delivery rate while ensuring good energy consumption. The approach is also able to tolerate and correct nodes failure problems. The authors in [129] presented a cross-layer MAC protocol suitable for homogeneous and heterogeneous wireless networks. The scheduling relied on a unique structure of packets that efficiently employs routing information to transmit multiple data packets through multiple multi-hop flows. The approach dynamically optimizes the scheduling mechanism and uses a FSP control (Flow Setup Packet) to manage the multi-stream traffic and reduce the latency.

Chou et al. proposed in [130] a centralized slot allocation algorithm based on the routing tree information called CoLaNet. This approach uses a two-hop vertices coloring algorithm. The algorithm gives color to nodes that have the maximum degree in the routing tree first, so that each node is not given a color taken by its one-hop or two-hop neighbors. After the execution of the algorithm, the colors are transformed into slots and the total number of slots represents the schedule length.

Table 2.7 outlines a brief comparison of the different cross-layer approaches presented in this section.

2.5/ MAC AND ROUTING PROTOCOLS ISSUES

From the literature, we observe some existing shortage when considering the MAC and routing approaches. For instance, when considering cross-layering approaches, the focus is mainly devoted for solving energy efficiency problems to maximize the network lifetime. Most of these approaches are centralized and do not essentially optimize the latency or neglect it completely by targeting other metrics, which is not suitable for critical

Table 2.7: Comparison of the cross-layer approaches

Algorithm	**Topology**	**Category**	**Routing structure**	**Objective**	**Limites**
ORW [125]	Distributed	Duty cycling	Opportunistic	Ennergy efficiency	Iiterative computation, require a global view of network
DRAND [126]	Distributed	Cont. free	Tree	TDMA Scheduling	Random slot allocation, high collision rate
ET-DRAND [70]	Distributed	Cont. free	Tree	Ennergy, scheduling	Increase communication latency
IDeg-ReLO [127]	Centralized	Cont. free	Tree	Latency	Central point of failure
CL-MAC [129]	Distributed	Cont. based	Any	Latency	Energy ignorance
Areacast [128]	Centralized	Cont. based	Any	Energy, overhead	Increase latency
CoLaNet [130]	Centralized	Cont. free	tree	Energy, scheduling	Increase latency

application.

In IoT networks, a communication protocol needs to be distributed. Indeed, with a large number of connected devices, managing the channel access with a central coordinator may obstruct the network performances. Therefore, nodes need to self-organize themselves to enable reliable communications. Moreover, these protocols should be dynamic in order to adapt with the topology variations caused by the wireless nature of the IoT network. Furthermore, MAC and routing protocol design should consider the broadcast nature of the wireless channel by minimizing collision and interference issues. In addition, as the radio spectrum and the available bandwidth are limited, the communication protocol should ensure an optimal usage of the available resources. Besides, these protocols need to address the energy efficient issue in the long-term (which is poorly addressed), particularly when developed for specific applications like data gathering and monitoring.

2.6/ Conclusion

This chapter presented a review of the MAC and hierarchical clustering approaches in wireless networks. In this regard, we presented a background of the Mac issues and classification, then we exhibit a taxonomy of the routing protocols, more precisely, we addressed the hierarchical clustering classification. Afterward, we approached the cross layering concept between the MAC and the network layer and we addressed the MAC-aware routing approaches. This study allowed us to take advantage of the deficiencies of the proposed protocols in the literature to introduce new clustering and cross layer approaches for large scale IoT networks. The proposed approaches aim to enhance the conventional wireless networks to structure the network topology and enable an effective management of the communication channels. The design and evaluation of these approaches is the subject of the following chapter of this thesis.

II

CONTRIBUTIONS

The objective of our work is to provide efficient management of the communication channels and the topology in large scale IoT network. We adopt different approaches within the two sub layers directly involved in the network communication, notably, the network and the MAC layers. In one hand, we considered a cross layer architecture to improve and manage the access to the shared communication channel. On the other hand, at the network layer, we addressed hierarchical clustering, which allows the management of the network structure in order to maintain connectivity of the large scale topology and improve the network performances.

In the first part of this chapter, we will present the first contribution concerning distributed TDMA scheduling. The aim is to improve communication latency by considering the information obtained from the routing layer via cross layer. In the second part, we present two clustering topology approaches. The first one focuses on the k-hop intra-clustering and uses the two-hop neighbor's connectivity to strengthen the cluster's connectivity. The aim is to optimize the cluster heads set to minimize the number of long range communication channels and expand the network lifetime. The second approach handles a distributed unequal clustering technique that uses the Log-Normal Shadowing model within an urban environment to evaluate the error rate of transmission within nodes and improve energy efficiency. This latter integrates a realistic communication model in unequal clustering to alleviate the hot-spot energy problem. In the third part, we focus on long-term energy optimization, and we present two approaches for dynamic clustering, considering the state of devices' health batteries and their degradation level. The proposed schemes efficiently manage the energy resource to enhance the battery behavior that extends the network lifespan in the long-term. The first approach presents a new dynamic clustering technique using an original long-term vision of energy optimization for IoT network. This approach considers the rechargeable battery aging behavior to extend the network life-time in the long term. The second proposal presents a novel SOH prediction based Clustering for Long-term Energy optimization. This latter uses an original metric to predict the forthcoming SOH of the devices' battery before the clusters formation, which significantly extends the system life. All proposals performance are evaluated by simulation using the same metrics with different network configurations (parameters, topology).

3

Latency improvement aware distributed cross layering

3.1/ Motivation

One of the main problems facing the development of MAC protocols is the dynamic topology due to the sensors' breakdown caused by exhaustion of energy resources or other factors like signal interference and nodes' mobility. These factors bring new challenges [22], involving communication channel allocation, which is of great importance. Efficient communication channel management optimizes the latency and minimizes interference and collision problems that affect energy consumption. One of the techniques used to reasonably manage communication channel is the allocation of time-slots to network nodes.

Typically, conventional slots scheduling algorithms are centralized, where the BS requires interfering information from every network device. The resulting schedule is close to the optimal solution [70]; however, it is not appropriate for dynamic networks since the slot schedule is recomputed after each topology alteration when a node joins or leaves the network. Moreover, these approaches face the major problem of the central point of failure hence the slots schedule cannot be performed. To adapt to dynamic topology of the network, slots allocation algorithms should have topology adaptability that effectively affect the available transmission slots according to the requirements. That is why the distributed techniques are more appropriate.

Traditional distributed protocols assume that the MAC and Routing layers act independently, which affects the latency metric. In this regard, the contribution presented in this chapter addresses the cross layer contention-free scheduling algorithms. More precisely, we focus on the problem of developing a weight-based TDMA-MAC scheduling algorithm in a distributed fashion. The proposed solution is a MAC aware routing protocol that takes advantage of routing information to improve the communication latency and the transmission schedule length.

To clarify the idea, the following example exhibits an unsuitable MAC decisions situation regarding the routing information. The network is modeled into a graph of 6 nodes, where the vertices and the edges represent the network devices and the communication links among devices respectively, as shown in the left side of the figure 3.1. One possible routing tree for this network is given on the right side of the figure.

In figure 3.2, nodes TDMA schedule is represented by vectors where each vector represents the time slot schedule used by the current node for: sending (marked in bold letter), receiving data (marked in italic letter), or idle (empty slot). This example aims to illustrate that the shortest path built by the routing protocol (network layer) does not always ensure the shortest communication delay due to the scheduling decision made by the MAC layer. In figure 3.2 (a), the TDMA schedule is generated without considering the routing information. For instance, the average communication latency to relay data packet from the source node D to the reception node A is 8 time slots, i.e. 1 time slots to send the packet from D to C, 4 slots to send the packet from C to B and 3 slots so that node B sends information to node A. Whereas in figure 3.2 (b), the scheduling is performed according to the routing information (by using cross layering). Therefore, node C is scheduled straight after its routing child nodes E and D, B is scheduled after its children C and F, and A is scheduled right after its child B. With this scenario, the average latency generated by the new schedule is reduced to 5 time slots. Indeed, 5 time slot are required to relay data packet from source node D to A, i.e. 1 time slot to send information from D to C, 2 slots so that node C relay the data packet to node B and 2 slots so that node B sends information to node A.

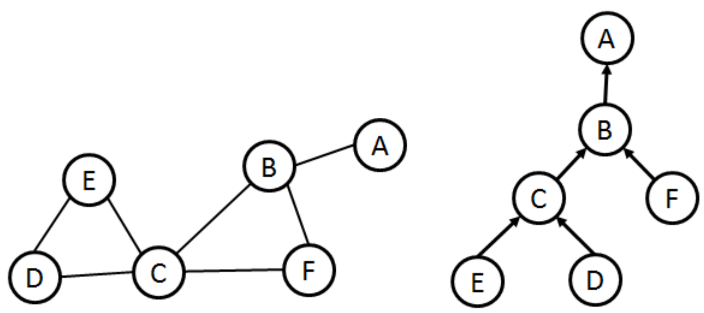
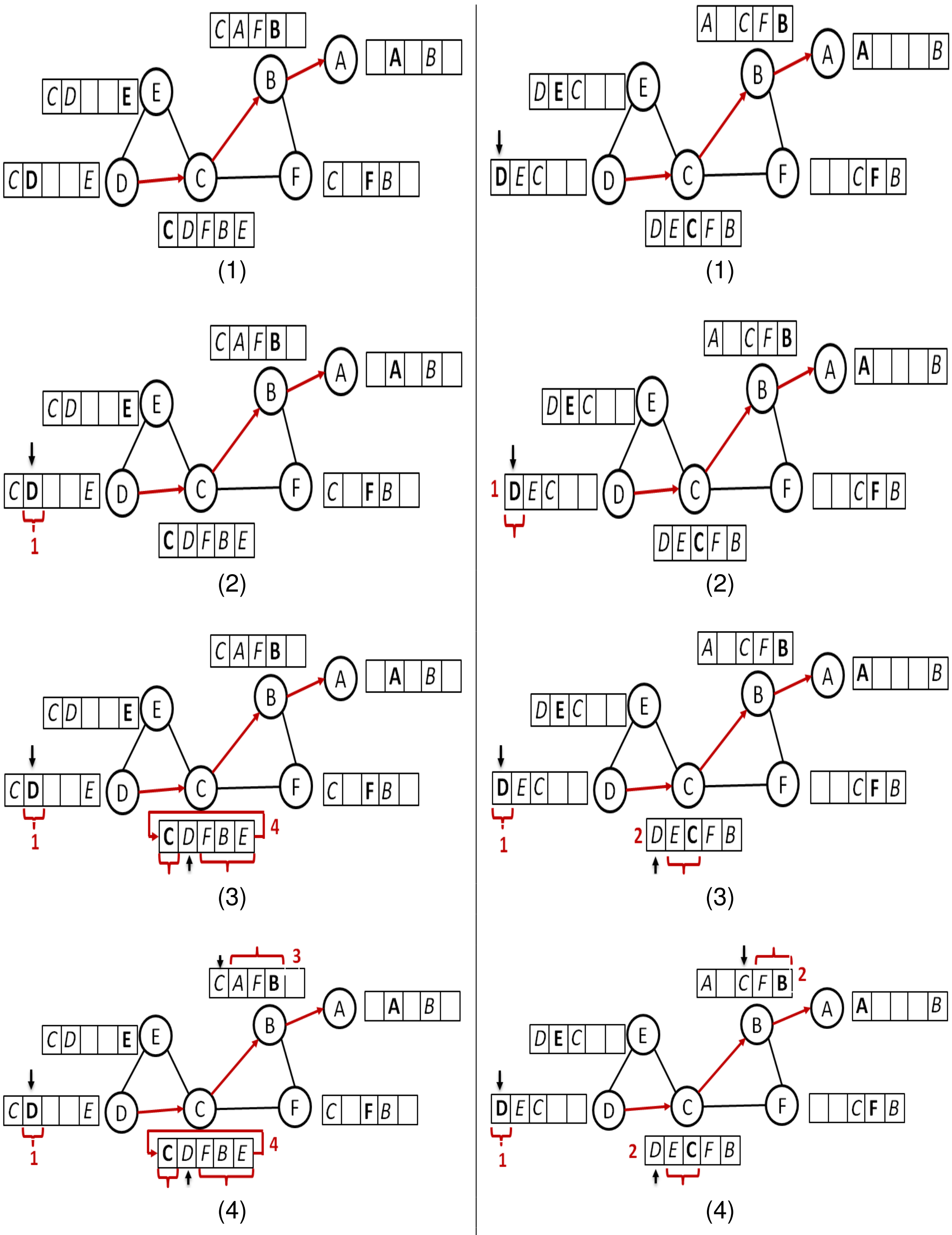


Figure 3.1: An example of a wireless network modeled as a graph (left) and its routing tree (right).

Therefore, the TDMA can take advantage of the spatial information coming from the routing layer to optimize the communication latency.



(a) TDMA Without cross layering

(b) Cross layer based TDMA

Figure 3.2: Two TDMA schedules of network nodes in Fig. 3.1

3.2/ CONTRIBUTION

This contribution presents a new Weight-Based Distributed TDMA Scheduling Algorithm (WB-DTSA), a distributed MAC-TDMA scheduling approach for IoT and sensor networks. The idea is to use the cross layering technique between the Routing and MAC layer to reduce the energy consumption and the transmission latency, which is suitable for critical applications. Since each node knows its interfering degree and its parent in the routing tree T, the idea is to take advantage of the routing tree information and give the transmission slots to nodes based on their position in the routing tree. A bottom up traversal is applied on T, while privileging nodes with long routes toward the BS (node with high tree level) by allocating them earlier slots in the transmitting schedule. In this contribution, each node in T performs a local slots schedule for its children, i.e. parents at level i in the routing tree allocate slots to their children at level $i+1$. Each node, except the BS, gathers the interfering information from its neighborhood and forwards them to its parent to generate the local schedule. Interfering nodes with different parent may receive similar transmitting slots. Thus, to avoid this scenario, we adopt a weight-based approach during the slot allocation process. The approach gives priority of nodes with higher weight, i.e. the node with higher weight obtains the slot first. Moreover, nodes with many neighbors are more challenging to schedule on existing slots than those with fewer neighbors. Therefore, we propose to combine two metrics to perform the weight constraints: the node interfering degree and the node average distance.

- **Node Interfering Degree (*IDeg*)**: This criterion denotes the total number of neighbors. This metric is used during the scheduling to prioritize nodes with large interference range. This metric is calculated by in equation 3.1.

$$N_{IDeg}(i) = \{\ j \in V \mid \exists (i, j) \in E \wedge P_i \neq (P_j \vee j)\ \}$$

$$IDeg_i = |N_{IDeg}(i)| \tag{3.1}$$

- **The Average Distance (*AD*)**: It stands for the average distance between a node and its neighborhood. The AD metric of a node i is calculated using equation 3.2.

$$N(i) = \{j \in V \mid (i, j) \in E\}$$

$$AD_i = \frac{\sum_{j=1}^{|N(i)|} (Dist(i, j) \mid (i, j) \in E)}{|N(i)|} \ \wedge\ j \in N(i) \tag{3.2}$$

The weight W_i of node i is calculated based on the previous parameters using equation 3.3. α and β are the weighting coefficients for the corresponding system metric. Where $\alpha + \beta = 1$

$$W_i = \alpha * IDeg_i + \beta * AD_i \quad (3.3)$$

Each weighting coefficient expresses the importance of the related metric relative to the others. In this work, both $IDeg$ and AD are perceived as important metrics and after performing several experiment scenarios, we assigned them an equivalent weighting coefficients which renders the best result. We choose $\alpha = 0.5$ and $\beta = 1 - \alpha = 0.5$ as weighting coefficients values for the experiments in this study. These weighting coefficients can be adjusted according to the system requirements. WB-DTSA runs in rounds. A round is measured in slots and the number of slots used in one round is adapted dynamically depending on the estimations of network delays. Each node i maintains two lists: $Interfering\_slots\_list_i$, containing the interfering neighbors slots and $Request\_list_i$ representing the list of children that have sent their interfering information to i and wait for the local schedule. A node can be in one of the three states: IDLE, REQUESTER, SCHEDULER. Figure 3.3 presents the transition state diagram of WB-DTSA.

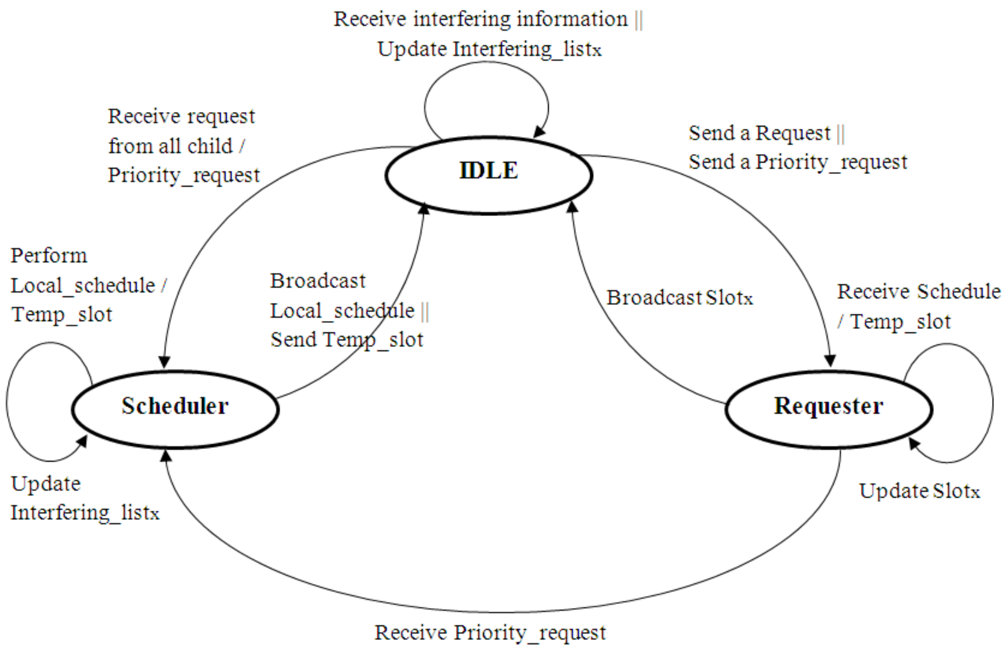


Figure 3.3: State diagram of WB-DTSA.

The approach works as follow.

1. First, each node i measures its W_i value and broadcasts this value to all its neighbors $N(i)$.
2. To obtain a slot, i collects interfering information from its neighborhood. Then:
   – If i is a leaf, it adds the collected interfering information to the slot request message ($Request_i$).

- Otherwise, i adds the collected interfering information and its children slots information to $Request_i$.

i sends $Request_i$ to its Parent P_i.

3. When P_i receives a request message from i, it adds i children slots information to its $Interfering\_slots\_list$.

4. When P_i receives request messages from all its children:

   (a) P_i sorts it $Request\_list$ according to requester weight, so that the node i having the highest weight will be scheduled first.

   (b) Find the minimum slot for i, while taking into consideration its interfering list.

   (c) Repeat step (b) for all $Request\_list$ elements.

   (d) After the slot allocation process, P_i broadcasts the resulting local schedule to its children.

5. When i receives the local schedule. It updates $Slot_i$, then broadcasts this slot to its neighbors, so it can not be used by interfering neighbors with lower weight.

When using only one type of request messages, a deadlock scenario may occur during the scheduling, as depicted in Fig. 3.4. In this scenario, node y waits for interfering information from the weighted node z (since it needs interference information from this neighbor to be able to send a slot request to its parent node x). Same case for z and x, where z is waiting for the interference information of its neighbors x. This latter (node x) is waiting for its son (node y) to send him this interference information. In this case we observe that a mutual waiting loop is formed between nodes y, z, x (eg. $y \xrightarrow{\text{wait for}} z \xrightarrow{\text{wait for}} x \xrightarrow{\text{wait for}} y$) as illustrated (the red cycle) in Fig. 3.4, which causes a deadlock situation that obstruct the algorithm execution. To bypass this deadlock scheme, we use Priority request messages ($Prio\_request_i$), which contains the interfering neighbors slots and $Request\_list_i$ representing the list of children that have sent their interfering information to i and wait for the local schedule as follows:

(a) When a node i finds that it has the highest weight, it sends a priority request message $Prio\_request_i$ to P_i.

(b) After that P_i receives $Prio\_request_i$, P_i searches for the minimum temporary slot available $tempo\_slot_i$ and sends it directly to i.

(c) When i receives a $tempo\_slot_i$ message, it updates $Slot_i$, then broadcasts this slot to its neighbors.

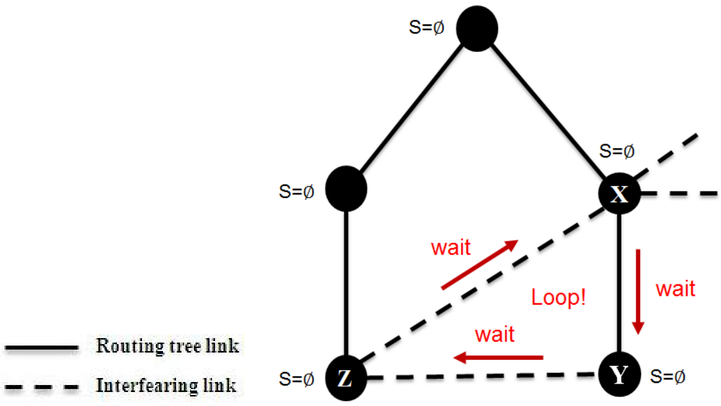


Figure 3.4: Deadlock example.

Figure 3.5 illustrate an execution example of WB-DTSA slots scheduling process in a small network (modeled by a graph). Nodes in black represent the nodes with assigned slot and the white nodes represent the nodes that are still waiting for a transmission slot. Figure 3.5 (a) and 3.5 (g) present, respectively, the initial and the final network state where all nodes have reserved there slots. In this example, each node collects the interfering slots assigned to its neighborhood and transmits them to its parent node in the routing tree. The parent performs a local slots schedule for its children by considering their interfering slots. Accordingly, a bottom up allocation of the transmitting slots is performed privileging nodes with long routes toward the BS, starting with the leaf nodes by allocating them earlier transmitting slots. This reasoning will be followed, by recurrence, by all network nodes. In the final network state, we obtain a TDMA schedule in which the transmission slots of the child nodes are close to that of their parents in the routing tree, which will reduce the latency of data aggregation in the network. The pseudo code of WB-DTSA is shown as algorithm 1. This approach takes into account messages generated by the routing layer other than those generated by the MAC layer. It may increase the number of messages generated by the algorithm during the execution and result in an extra amount of energy consumed. However, to avoid this disadvantage, network clustering techniques are used.

3.3/ Network model

Wireless networks are generally represented by a graph $G = (V, E)$, where V is the set of nodes (devices or sensors) and E is the set of edges (communication or interfering links). $N(i)$ denotes the set of the direct neighbors of node $i \in V$. $Deg(i) = |N(i)|$ represents the number of neighbors (degree) of i and $SLOT_i$ is the transmitting slot allocated to i. $Dist(i, j)$ is the distance between i and j. We use the UDG (Unit Disk Graph [131])

Algorithm 1 WB-Distributed TDMA Scheduling Algorithm

Variables:
W_i: the weight of i.
P_i: parent of i in the routing tree T.
$Interfering\_slots$: list of interfering neighbors slots.
$Request\_list_i$: list of slot requesters.
$State_i$: indicates the state of i.
$Leaf_i$: indicates if i is a leaf in T.
$Count\_child\_request_i$: counter of child slot request messages received.
Wn_i: neighbors of i that have higher weight.
$Child\_slots\_list_i$: children slots list of i in T.
Output:
$SLOT_i$: the transmiting slot of i.

if $State_i = IDLE \wedge Wn_i \in Interfering\_slots_i$ **then**
    if $!Leaf_i \wedge |Child\_slots\_list_i| = |Child_i|$ **then**
        Update $(Child\_slots\_list_i, Interfering\_slots_i)$
    Send $(Request_i, Interfering\_slots_i)$ to P_i
    $State_i = Requester$
Upon receiving $Request_j$ message:
    Add j to $Request\_list_i$
    Add $Child\_slots\_list_j$ to $Interfering\_slots$
    $Count\_child\_request_i$++
    if $Count\_child\_request_i = |Child_i|$ **then**
        $State_i = Scheduler$
        $Sort(Request\_list)$
        $Schedule = Allocate\_Slots(Request\_list)$
        Broadcast($Schedule$)
        $Clear(Request\_list_i)$
        $Count\_child\_request_i = 0$
        $State_i = IDLE$
if $\nexists j \in N_i : W_j > W_i$ **then**
    Send $Priority\_request_i$ to P_i
    $State_i = Requester$
Upon receiving $Priority\_request_j$ message:
    $State_i = Scheduler$
    if $Slot_j \neq \emptyset$
        Delete $Slot_j$ from $Interfering\_slots_i$
    $Count\_child\_request_i$++
    $Temporary\_Slot_j = Find\_min\_slot(Interfering\_slots_j)$
    Add $Temporary\_Slot_j$ to $Interfering\_slots_i$
    Send $Temporary\_Slot_j$ to j
Upon receiving $Schedule \vee Temporary\_Slot_i$ from P_i:
        Update $Slot_i$
        Broadcast $Slot_i$
        $State_i = IDLE$
Upon receiving $Slot_j$ from $j \in N(i)$:
    Add $Slot_j$ to $Interfering\_slots_i$
    if $Slot_i = Slot_y$
        $Slot_i = \emptyset$
        $State_i = IDLE$

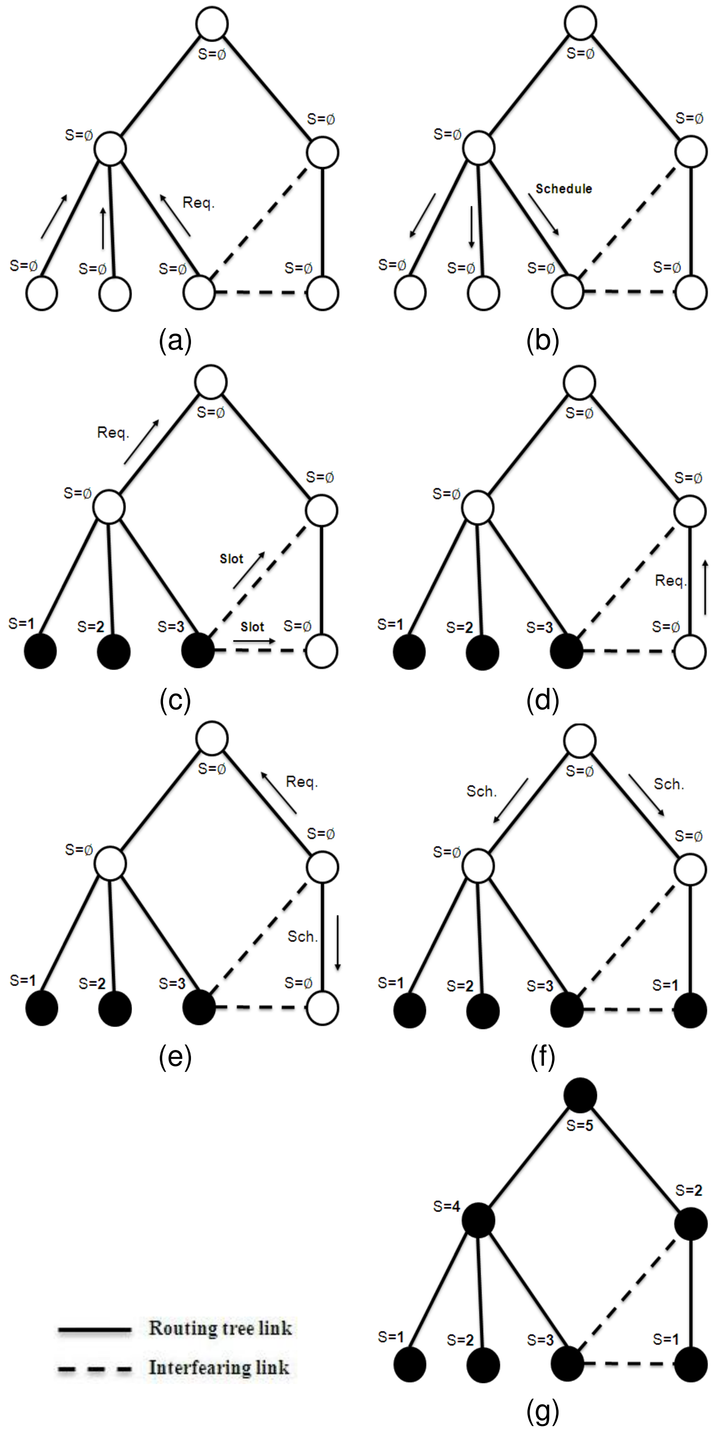


Figure 3.5: Execution scenario of WB-DTSA slots scheduling process.

connectivity model, in this model a link exists between two nodes i and j if they are within each other's transmitting range R ($Dist(i, j) \leq R_j \wedge Dist(j, i) \leq R_i$). We consider that communication links are symmetric, i.e. all devices have the same transmitting range. The routing tree T used for data aggregation toward the base station (BS) is formed by a subset of communication links E_t. We denote by P_i the parent of i in T. A link that does not belong to T $(e \notin E_t)$ is considered an interfering link because messages transmitted via this link do not reach their destination. Network devices use a single channel transceiver, i.e. they cannot receive from multiple senders or send and receive at the same time. Any node in T (except the BS) can transmit a packet to the BS through the routing tree path.

3.4/ Energy Model and Transmission Reliability

Wireless network devices perform diverse functionalities, such as sensing, processing, transmitting and receiving information. Among all those functionalities, wireless communication is the one that consumes the major portion of the energy and is far greater compared to the energy dedicated to computation [32]. Indeed, transmitting one bit over a distance of 100 meter requires approximately the same energy as executing 3000 instructions [122]. The energy model for this work is inspired from the model used in [19; 123; 90]. In this model, the transmitter wastes some energy to run the radio electronics and the power amplifier. On the other side, the receiver dissipates an amount of energy to run the radio electronics, i.e. detect and decode the radio signal as shown in figure 3.6.

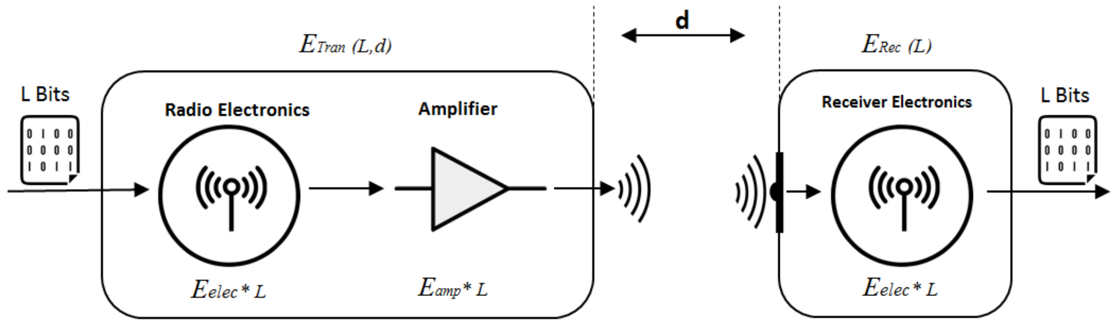


Figure 3.6: Radio energy dissipation model

The energy consumed by a node for transmitting an l bit data packet over a distance d is the sum of the energy consumption of the radio electronics circuit and the power amplifier circuit. It is measured as in eqation 3.4.

$$E_{Tran}(l, d) = E_{elec} \times l + E_{amp} \times l \tag{3.4}$$

E_{elec} depicts the energy dissipated per bit to operate the transmitter or the receiver circuit. The amount of energy consumption depends on different parameters, such as digital coding, modulation, filtering and spreading of the signal. Whereas, E_{amp} is the power amplification coefficient, it relies upon the distance to the receiver and the acceptable bit-error rate. E_{amp} is formulated as in equation 3.5.

$$E_{amp} = \begin{cases} \varepsilon_{FS} \times d^2 : d < d_0 \\ \varepsilon_{MFS} \times d^4 : d \geq d_0 \end{cases} \tag{3.5}$$

Where:

$$d_0 = \sqrt{\varepsilon_{FS} / \varepsilon_{MFS}} \tag{3.6}$$

On the other side, the energy required for the reception of a message of size l bits is given by equation 3.7.

$$E_R = E_{elec} \times l \tag{3.7}$$

In this model, the energy consumption by a transmitter is proportional to the threshold transmission distance d_0. According to the distance value d_0 (Equation 3.6), the propagation loss can be modeled as a free-space model or a multi-path attenuation model. ε_{FS} and ε_{MFS} are used for the free space and multi-path model respectively, they represent the characteristics of the transmitter amplifier which depend on the required receiver sensitivity and receiver noise [132].

3.5/ Simulation configuration

The parameters employed for evaluating the proposed algorithm are described in this section. The performances of the approaches are assessed via simulation. We used a simulator implemented in Java based on using Java Universal Network/Graph (JUNG) [133], a Java based library that allows the analysis, visualization, and modeling of networks as graphs. We assume that network nodes have the same transmitting range $R = 30\ m$ using the UDG connectivity model (Unit Disk Graph [131]). The network is made of a variable number of nodes deployed over a square area $\Delta^2 = 200 * 200\ m^2$. The BS is positioned in the upper left corner of the deployment area. The initial energy of each node is equal to 1 joule, and the data packet size is 100 bytes. The density δ represents the average number of neighbors in the transmitting range of a node. It is stated as $\delta = \pi * R^2 * n/\Delta^2$.

The network density may vary depending on the deployment environment and the application requirement. A high density may ensure good network connectivity and guarantee superior redundancy for better reliability of the aggregated data. In contrast, a high density increases the communication interference and data collision rate, which raises energy consumption. Therefore, the number of nodes deployed in the simulation area changes to obtain different densities. For this experiment $\delta \in [4, 50]$, to evaluate the performance of our algorithm, we use the following metrics:

- **The average energy consumed**: It denotes the average energy consumed per node before obtaining a transmitting slot during the time scheduling process.
- **The average latency**: the average end to end delay for each node (except the *BS*). The end to end delay for a node represents the number of time slots needed for a generated packet to reach the *BS* through multi-hop communication.
- **The schedule size**: the number of slots used by the algorithm to generate the TDMA schedule (the size of the TDMA schedule).

3.6/ Performance evaluation

The performances of the proposed approach are compared to Drand, E-T-Drand and I-Drand according to the metrics described in section 3.5. Figure 3.7, 3.8 and 3.9 show the obtained results. Multiple scheduling schemes have been proposed in the literature. We compare our approach to the classic Drand [126]. E-T-Drand [70] and I-Drand [11].

Figure 3.7 shows the average energy consumed by a node to procure a transmitting slot according to different values of density. Usually, as the density increases the energy exhausted by the four algorithms expands as well. When the density is lower than 8 ($\delta \leq 8$), the energy consumed by Drand surpasses that of WB-DTSA. This is due to the low network connectivity that increases the number of failure messages generated in Drand. The energy depleted by I-Drand and E-T-Drand increases slowly because these approaches use priorities during the time allocation. E-T-Drand is based on the energy information of the node and its neighbors, which reduces the number of messages generated. The average improvement of E-T-Drand is 33.9% compared to Drand. I-Drand uses the routing layer information to reduce the number of collision messages and the per-link latency, hence, it shows an average improvement of 18.4% compared to Drand. Unlike Drand and E-T-Drand which only consider MAC layer messages. WB-DTSA takes into account the messages processed by the routing layer in addition to those generated by the MAC layer. This considerably increases the number of messages processed by the approach during the execution and expands the amount of energy consumed. Although the energy consumed by our approach is larger than Drand (19.8% more), it shows superior performance in terms of latency and schedule length efficiency as shown in figures 3.8 and 3.9.

Figure 3.8 illustrates the average latency of each node with different node density values. Compared to the other approaches, our approach significantly diminishes the communication delay and generates a smaller latency. Compared to I-Drand, our approach show a latency amelioration of 38% and improves the global latency by 47.1% compared to Drand. Indeed, Nodes are scheduled based on the order given by a bottom-up traversal

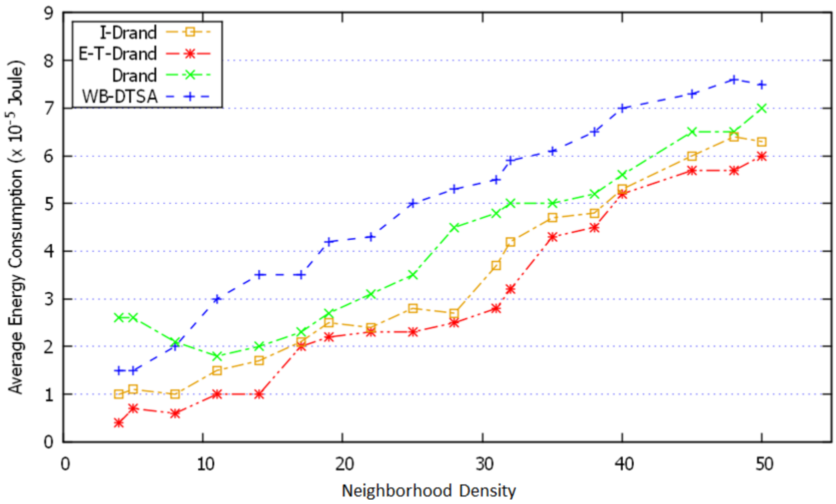


Figure 3.7: Average energy consumed

of the routing tree, privileging nodes with long routes toward the BS by trying to assign them earlier slots in the schedule. As a result, the per-link latency is reduced, and the global end-to-end delay is improved, which is appropriate for critical applications.

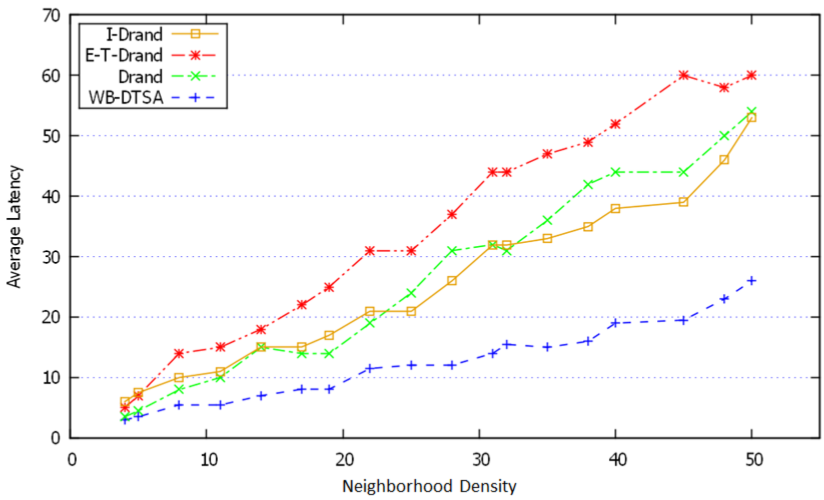


Figure 3.8: Average communication latency

Figure 3.9 shows the average number of transmitting slots used in the scheduling. We notice that the schedule size depends on the network density. Moreover, the arrangement of slot allocation priority following a bottom-up traversal reduces the schedule length. Figure 3.8 shows that the schedule length of E-T-Drand and Drand is nearly equal, while WB-DTSA and I-Drand both reduce the number of time slots used. Our approach improves the schedule length by an average of 22.8%, which denote that our proposal significantly reduces the number of time slots used in the scheduling.

Table 3.1 illustrate a comparison of the different algorithms to the classic Drand [126].

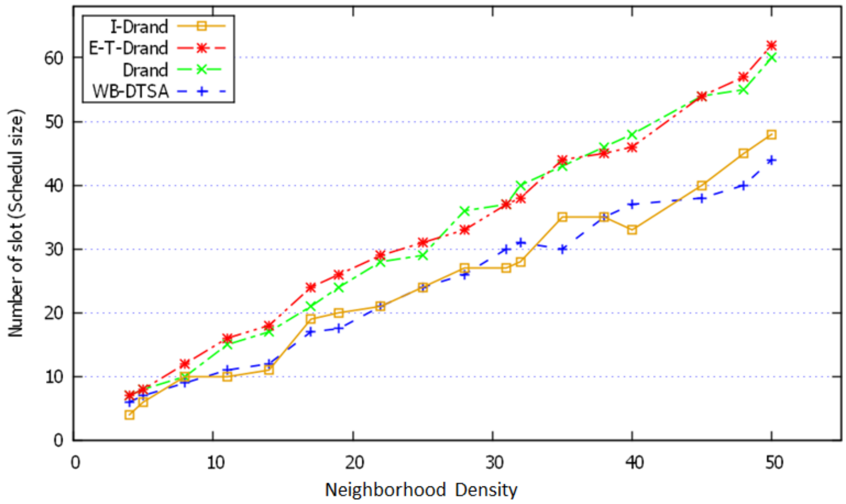


Figure 3.9: Average number of time slots used in the scheduling

It synthesizes the performance gain and loss, considering different density values with different metrics.

Table 3.1: Performance gain and loss of WB-DTSA, E-T-Drand and I-Drand compared with Drand

Algorithm/Density		**[4-20[**	**[20-35[**	**[35-50]**
WB-DTSA	***Latency***	35.9%	53.8%	54.5%
	Schedule size	19.8%	23.6%	26.6%
	Energy consumption	-20.8%	-23.5%	-15.3%
E-T-Drand	***Latency***	-17%	-36.1%	-19.4%
	Schedule size	3.7%	-4.5%	-1.8%
	Energy consumption	39.7%	29%	11.5%
I-Drand	***Latency***	6.7%	3.6%	16.8%
	Schedule size	27.9%	24.2%	20.2%
	Energy consumption	30.3%	19%	6.1%

3.7/ CONCLUSION

In this chapter, we presented a new distributed cross-layer TDMA scheduling algorithm named WB-DTSA. In contrast to the typical centralized scheduling approaches, the proposed contribution is distributed to tolerate the dynamic variation of the topology and avoid the central point failure problem. The novelty is that the proposal takes advantage of the routing tree information and exploits the knowledge of the neighborhood to perform the communication slot scheduling. Indeed, the transmitting slots are attributed to nodes based on their position in the routing tree while privileging nodes with long routes toward the base station by allocating them earlier transmitting slots to reduce the latency. The proposal employs a weight based technique that considers the average distance and the interference degree among nodes and their neighborhoods to efficiently improve the slot allocation. Several simulation experiments have been executed to estimate the performance of the approach. Simulation results have shown that the proposal has better performance in comparison to similar approaches in terms of communication latency and schedule size.

4

Distributed self-stabilizing approach for the k-hop intra-clustering aware neighbors topology knowledge

4.1/ Motivation

Various clustering protocols concentrate on the multi-hop inter-clustering [134; 135; 136] between the cluster heads (CHs) and the base station to increase the network lifetime. However, only a few consider the intra-cluster communication (between devices and their CH). Existing intra-clustering schemes [110; 137] usually presume a direct connection between Cluster Members (CMs) and their CHs. Thereby, a great number of clusters are created. In a wide scale network, the distance between nodes and their CHs may not be short enough for communication. Hence, direct communication becomes obstructive, and k-hop intra-clustering communication should be employed to allow network scalability. The k-hop intra-clustering model perceived a substantial concern from the research community [138; 117], as it exhibits (in a better way) the features of a well-organized network. Indeed, k-hops intra-clustering model restrains the number of clusters and, consequently, reduces the number of long-range communication channels and extends network durability [136]. The objective of k-hop clustering is to arrange nodes into clusters where the path between CMs and their corresponding cluster head is shorter than k hops of distance. It offers a robust topology in dynamic networks since cluster members may not be in direct liaison with their CH. Hence reconfiguration processes (cluster head election and re-affiliations) are constrained [29].

K-hop intra-clustering has proved many advantages to alleviating the congestion problems and prolonging the network lifetime [139; 140]. However, most of the k-hop intra-clustering protocols are centralized and only consider the node's direct neighbors' information, which lacks robustness. Indeed, with a centralized k-hop scheme, the BS is a central point of failure, i.e. a potential BS failure will collapse the execution of the whole protocol. A loss of a critical node or a communication error could subsequently cause a severe clustering failure because some data are usually of crucial importance in a centralized approach. Moreover, with the network scaling, BS may become a performance bottleneck. In this context, this paper presents a new distributed k-hop intra-clustering protocol that considers the two-hop neighbors' topology knowledge. The aim is to optimize the set of CHs to reduce the number of long-range communication channels to avoid congestion issues and extend the network lifetime.

The main contributions of this work are as follows:

- A new connectivity metric is introduced to elect the set of proper CHs. The novelty of this metric is taking into account the two-hop topology of the current node, and its surrounding neighborhood (instead of using the traditional direct neighbors connectivity), in order to strengthen the cluster's stability.
- The design of the algorithm is based on a distributed self-stabilizing system. We prove that the algorithm stabilizes, and converges within $O(n + k)$ rounds, which represents the upper bound of the time complexity, n is the number of network nodes and k is the depth threshold of the clusters. This perspective enables network devices to efficiently tolerate possible failures that can occur locally in the dynamic topology.
- The proposed approach produces clusters with an energy efficient topology by reducing the distance between nodes and their respective CH. The approach is particular in that it constructs the intra-cluster links in a distributed fashion rather than using a centralized algorithm executed by the CH.

4.2/ Network Model and Algorithm Objective

In this contribution, we adopted the network model described in the previous chapter (section 3.3) by adding a new term used in this approach. For instance, $N_k(i)$ denotes the set of k-hop neighbors of the node $i \in V$, $N_{<k}(i)$ represents the set neighbors at distance $< k\text{-}hop$ from i.

The proposed solution objective is to optimize the number of CHs (e.g. reduces the number of CHs while maintaining the network connectivity), reduce the waste of energy caused by long-range communication channels, and decrease the risk of congestion. More precisely, we aim to find the smallest set C of selected cluster heads $\{CH_n \ (n \in$

$1, 2, ..., |C|)\}$ using a two-hop neighbors connectivity metric while respecting a particular maximum hops constraint k between each node and its cluster head. Considering that Mn is the set of nodes that belong to the cluster coordinated by the cluster head CH_n. The objective of the proposed mechanism is formulated as follows:

$$C = (Min\ |\bigcup CH_n \subseteq V|)\ \wedge\ (\ \forall i \in M_n,\ \forall\ CH_n \in V) :$$
$$Dist(i, CH_n)\ is\ Minimized$$

The communication over multi-hop short-range is generally more energy efficient than directly transmitting in a single-hop long-range communication [141]. Wide intra-cluster hops expand the coverage of the CHs, which reduces the number of elected CHs and the energy dissipation. However, in the case of a dense network, the interference rate will also rise, augmenting the waste of energy due to collision problems.

4.3/ Contribution

In this section, we present our distributed k-hop intra-clustering approach for wireless networks named Distributed Clustering aware 2-Hop Connectivity (DC2HC). The approach consists of grouping nodes with high connectivity into k-hop clusters. As cluster heads require more energy compared with other nodes, the number of CHs should be reduced. Therefore, DC2HC aims to produce an optimized number of clusters while preserving the network coverage, reducing the number of isolated nodes, and offering extended network durability. The algorithm approves a weight based mechanism during the cluster heads election process. This latter consists of selecting as a CH node with the maximum weight in its k-hop neighborhood. Multi-hop spanning trees [142] are shaped inside each cluster where CMs use the Received Signal Strength Indicator (RSSI) to select their parents in the routing path leading to the related CH. Hence, clusters will be constructed with a topology that spends less energy. Nodes weight combines three metrics:

takes into account 2-hop connectivity and pattern consumption. First the nodes must evaluate their node degree and their neighbor's node degree to calculate the Mean Connectivity Degree (MCD), which consists on the addition of its own degree, d i , and its neighbors' degree, d j , divided by its own degree plus one:

1. **Two-hop connectivity ratio (TCR)**: this parameter represents the connectivity ratio of a node relatively to its neighborhood. Each node computes the average connectivity within its two hops neighborhood ($N_{\leq 2}(i)$) using Φ_i, then compares the acquired value with the average two hops connectivity of its neighborhood. The TCR_i value of a given node i is estimated according to Equation 4.1. A higher TCR_i value sig-

nifies that i, besides of having a good cardinality, is well connected with its two-hop neighborhood and represents a convenient CH candidate to sustain the network connectivity as it covers a wider number of nodes within the maximum hop constraint k.

$$\Phi_i = \frac{\left[\sum_{j=1}^{|N_{\leq 2}(i)|} |N(j)| \ / \ j \in N_{\leq 2}(i)\right] + |N_{\leq 2}(i)|}{|N_{\leq 2}(i)| + 1}$$
$$TCR_i = |N_{\leq 2}(i)| - \Phi_i \tag{4.1}$$

Figure 4.1 illustrates a CHs election example in a small network using the two hop intra-clustering (modeled by a graph where the black nodes represent the CHs) using a conventional cardinality based Fig. 4.1 (a) metric in comparison with the TCR metric Fig 4.1 (b). In this example, as nodes A, B, C are potential CHs. With a conventional metric, as the three nodes have the same cardinality, either nodes A and C can be elected as CHs to regroup the network node. Consequently, two cluster will be formed in the network. Whereas, when using the TCR metric, the node B is well connected (compared to the other nodes) and will be selected as it has a better TCR value, which will allow the node B to cover more nodes within the same cluster. Thereby, the number of CH will be reduced while maintaining a good connectivity.

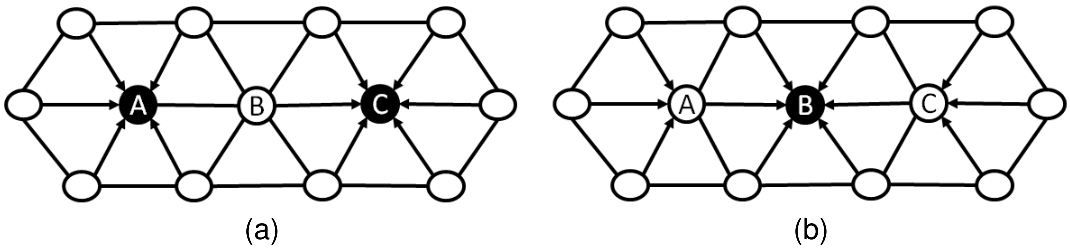


Figure 4.1: Example of CHs election using a conventional cardinality metric (a) versus the TCR metric (b)

2. **Residual energy (E^{ratio})**: the remaining energy of network nodes is introduced in the CH election process. The ratio of remaining energy of a node i is computed as in equation 4.2, where E_i^{init} is the initial energy of the current device and $E_i^{residual}$ is its remaining energy.

$$E_i^{ratio} = \frac{E_i^{residual}}{E_i^{init}} \tag{4.2}$$

3. **Communication reliability**: DC2HC uses Radio Signal Strength Indicator (RSSI) as a metric to quantify the quality of communications. The RSSI value (the received transmission power P_r) can be represented by the Log Distance Path Loss Model

[143] as in equation 4.3

$$P_r(d)(\mathrm{d}Bm) = P_t(\mathrm{d}Bm) - 10 \times \alpha \ log(d) - X_\sigma \tag{4.3}$$

P_t represents the power of the transmitter's radio signal in dBm. The distance d between the transmitter and the receiver is evaluated in meters. α is the path loss exponent that relies upon the environmental conditions ($\alpha = 2$ in the free space propagation model). X_σ is a Gaussian random variable used in the case of shadowing effect. Otherwise, it equals zero.

Therefore, the weight W_i of a node i, based on the previous parameters, is computed using the equation 4.4.

$$W_i = \alpha \times TCR_i + \beta \times E_i^{ratio} + \gamma \times P_r \tag{4.4}$$

where α, β, γ represent the weighted coefficients of the corresponding metrics with the constraint $\alpha + \beta + \gamma = 1$. The proposed scheme is designed to work under a typical network with various configurations to cover different use-case scenarios. Therefore, all the weight coefficients are considered equal.

In DC2HC, the clustering is completely distributed where each node has only a local perception of the network, which consists of its two-hop neighbor's knowledge. The algorithm is conceived by a set of rules of the form [If $condition$ then $action$], where the condition is a predicate specified over the node partial information. If the predicate is true, then the node is enabled to make a move (execute an action). This algorithm structure is inspired by the self-stabilization algorithms that are considered a valuable approach for tolerating transient faults in a network [144], which is preferable in an environment with a dynamic topology. We assume that each node i has a unique identifier and carry a set of variables that constitute the Local State Values of the node (LSV): node identifier (ID_i), its weight W_i, $Mych_i$(the relative CH of the current node), CH_{weight} (the weight of the relative CH), $Dist(i, Mych_i)$ that specify the distance to $Mych$ (in term of hops), $Deg(i)$, and its parent P_i in the shortest data aggregation path. The structure of LSV is illustrated in figure 4.2. Each node has a clustering record list CRL that contains a set of neighbor's information required by the clustering process. Nodes store the received clustering information (LSV beacons) in this list.

Table 4.1 summarizes various notations used in this contribution. DC2HC consists of three phases: the initialization phase, cluster heads election phase, and maintenance phase.

ID: 9	W: 10	Mych: 5	CH_w:10	Dist: k	Parent: 8	Deg: 20

Figure 4.2: Example of LSV structure.

Table 4.1: Notation used in DC2HC

Symbol	**Description**
ID_i	Identity of device i
W_i	Weight of node i (computed using Equation (4.4))
$Mych_i$	Relative Cluster Head of node i
CH_weight_i	Weight of $Mych_i$
$Dist(i, Mych_i)$	Distance between node i and $Mych_i$ (in term of hops)
P_i	Parent of i in the aggregation path toward the relative CH
Tr_i	Transmitting range of i
$Deg(i)$	Degree of node i (Number of node in the neighborhood of i)
LSV_i	Local State Variable of node i
CRL_i	Clustering Record List (Neighbors clustering information received)
TCR_i	Two-hop connectivity ration of node i
$Tx\_QUALITY$	Message used to integrate the network and measures the communication quality with the neighborhood
$CH\_Announc.$	Message broadcasted by nodes with a higher weight to inform their neighborhood about their current CH task

4.3.1/ Initialization Phase

The BS broadcasts a periodic beacon signal at the initial step. According to the received signal, nodes can estimate the transmission quality toward the BS (the RSSI value). After receiving this beacon, each node broadcasts a periodic $Tx\_QUALITY$ message in its transmitting range. After receiving a $Tx\_QUALITY$ message, a node updates its local state and replies by sending an LSV beacon. When the new node i receives LSV beacons from its neighborhood, it updates the CRL_i list and calculates its weight W_i using equation 4.4, then rebroadcasts a clustering message (LSV) that includes the updated information. The LSV beacon is broadcasted whenever local information change. Hence, network nodes can conserve coherent and updated information about their local environment. The initialization process is described in Algorithm 2. Figure 4.3 summarizes the initialization phase.

4.3.2/ Cluster Head Election Phase

The clustering process is described in Algorithm 3. We suppose no potential failure may appear while performing the CHs election phase. Initially, to elect a cluster head, a node

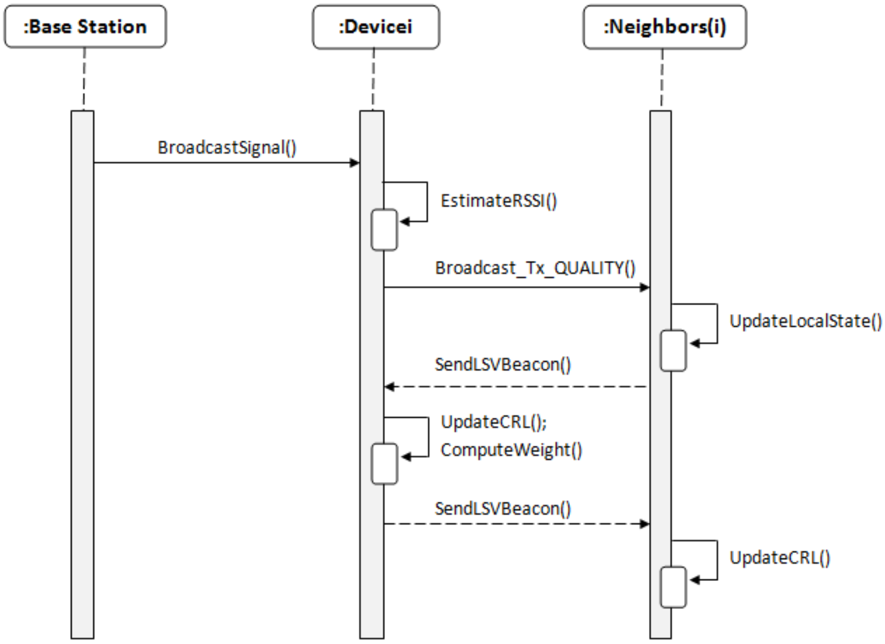


Figure 4.3: Sequence diagram of $DC2HC$ initialization phase.

Algorithm 2 $DC2HC$ Initialization phase.

Variables:
ID_i: identifier, $CRL_i = Null$, $N_i = \emptyset$
$Deg_i = 0$, $W_i = 0$, $RSSI_i = -100$ dBm

Upon receiving BS signal: Update $RSSI$ value
Broadcast $Tx\_QUALITY_i$ message
Foreach received LSV Beacon **do**
$|N_i| = |N_i| + 1$
Update TCR_i value using Equation (4.1)
$Deg(i) = |N(i)|$
Update CRL_i list
Endfor
Update W_i value using Equation (4.4)
Broadcast LSV_i Beacon
Upon receiving $Tx\_QUALITY_j$ message:
Send LSV_i beacon to j

i browses the list CRL_i and compares its weight with that of its neighbors and that of the CHs that dominate its neighbors (rule $R1$). If W_i is the greatest weight, i updates $Mych_i = ID_i$ (elect itself as a cluster head) and set $Dist(i, Mych_i) = 0$ (execute $R2$), then broadcasts a $CH\_Announcement$ beacon that includes the new CH information. Otherwise, i designates the node with the highest weight value in CRL_i list as its new CH. If two or more nodes have the same highest weight, the device with the highest ID is elected as CH. Rule $R1$ prevents the formation of poorly structured clusters; it enables network nodes to choose the most convenient cluster head in their k-hop neighborhood. Otherwise, the node selects the second weightiest node in CRL_i and so on until finding the weightiest CH with less than k hops of distance. When no CH meets these requirements, the current node elects itself as a new CH to avoid an isolated node situation. Node i updates $Mych_i$ with the identifier of the chosen node from (R1) and chooses the path with the minimum number of hops toward the CH according to the information gathered from its one hop neighborhood (using rule $(R3)$). Subsequently, nodes will be constrained to join the closest cluster that includes the weightiest CH, which avoids the formation of a height number of small clusters. The cluster heads election process is depicted in figure 4.4.

Algorithm 3 Cluster Head election phase.

Variables:
ID_i: identifier
$N(i)$: set of neighbors node of i
W_i: weight of the node i
Max_weight: temporary variable used to find the node with the highest weight
Output
CRL_i: clustering record list of i
$Mych_i$: the relative CH of i
$Dist(i, Mych_i)$: the minimal distance between i and its relative CH (in term of hop)

(R1): If $Max\_weight \neq Max(\{CRL[ID_j][W] \mid j \in (N(i) \vee Mych_{N(i)}) \wedge Dist(j, Mych_j) < k\} \bigcup \{W_i\})$ **then**

$Max\_weight = Max(\{CRL[ID_j][W] \mid j \in (N(i) \vee Mych_{N(i)}) \wedge Dist(j, Mych_j) < k\} \bigcup \{W_i\})$

$Mych_i = ID_j$

Else

If $i \neq j \wedge W_i = W_j$ **then**

If $ID_i \geq ID_j$

$Mych_i = ID_i$

Else

$Mych_i = ID_j$

(R2): If $(Mych_i = ID_i) \wedge CRL[Mych_i][Dist_i] \neq 0$ **then**

$Dist(i, Mych_i) = 0$
$CRL[Mych_i][Dist_i] = 0$
Broadcast $CH\_Announcement$ beacon

(R3): If $(Mych_i \neq ID_i) \wedge (Dist(i, Mych_i) \neq Min(Dist(j, Mych_j) \mid j \in N(i) \wedge Mych_i = Mych_j) + 1)$ **then**

$Dist(i, Mych_i) = Min(Dist(j, Mych_j) \mid j \in N(i) \wedge Mych_i = Mych_j) + 1$
$CRL[Mych_i][Dist_i] = Dist(i, Mych_i)$

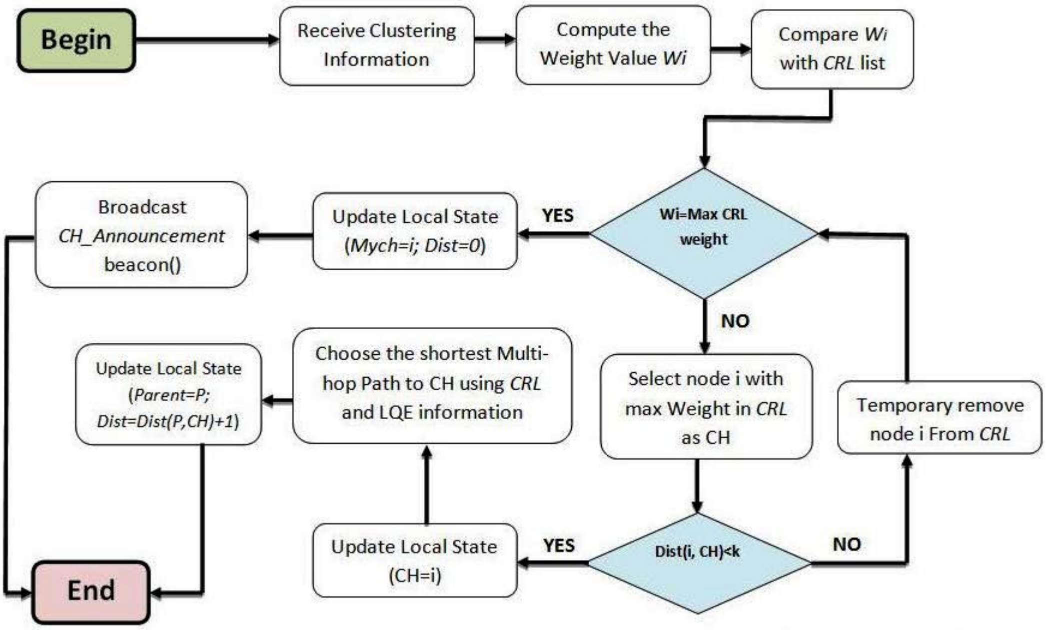


Figure 4.4: The cluster head election phase.

4.3.3/ Execution example

An execution scenario of the clustering process with $k = 2$ in a small network (modeled by a graph) is illustrated in figure 4.5. Green nodes depict the CHs, and the white nodes constitute the cluster members. Tables (b,d,f,h) indicate the different nodes parameters. Figure 4.5 (a) and (g) show, respectively, the initial network state and the final clustered network. Initially, network nodes received the signal message from the BS and exchanged their clustering information. At each stage, the variables of each node exploited during the clustering are illustrated in the corresponding table. Notice that in figure 4.5 (a), the weight of nodes 1 and 11 are the highest among all the nodes in their CRL lists. During the following round (Figure 4.5 (c)) these two nodes execute rule R1 to update their $Mych$ value (became CHs) and rule R2 to update the distance $Dist(i, Mych_i) = 0$. Next, nodes 1 and 11 transmit an announcement beacon to their neighbors. When the neighbors of these nodes receive the $CH\_Announcement$ message, they update their CRL lists and join the CH with the highest weight according to rule $R1$, then execute R3 to select the shortest path toward their CHs using the collected neighborhood information. At this point, two clusters are shaped in the network: $C1$ $\{CH : 1;\ CMs : 2, 3, 4, 5, 9\}$, $C2$ $\{CH : 11;\ CMs : 12, 13, 14\}$. Next (Figure 4.5 (e)), as new members joine the cluster, the members transmit an LSV message containing the updated cluster information to their surrounding neighbors. Hence, in the last state (Figure 4.5 (g)), the same scenario can be repeated for the rest of the nodes. The remaining unallocated nodes received the clustering information and updated their CRL lists. These nodes join the cluster with the weightiest CH, then update their distance using $R3$ $(Dist = 2)$ to form the final two hops cluster topology.

4.3.4/ Maintenance

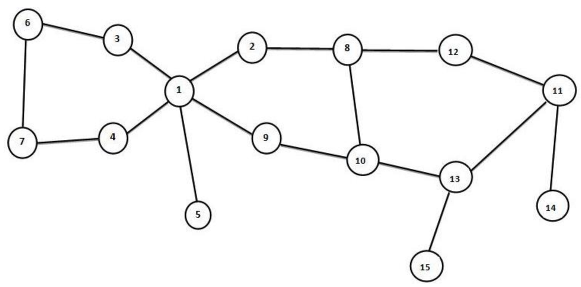
In a dynamic wireless network, recurrent topology changes can appear due to different factors such as device mobility, battery draining or lack of coverage, etc. The clustered structure must be resilient to an eventual node disconnections incident. Our approach assumes a continuous clustering maintenance, where devices regularly exchange control information. This mechanism is convenient for a dynamic scenario where disconnection events frequently arise [29]. The DC2HC re-clustering process can be locally driven and does not impact the entire network when disconnection events take place. Frequently updated control information enables a quick reaction to disconnection events. Thus, the clustered structure is more efficient and stable. The cluster maintenance phase is outlined as follows:

4.3.4.1/ Cluster Leaving

Each node monitors its surrounding environment through the exchanged beacon messages to keep track of the neighborhood members. When a node leaves the cluster (following a disconnection factor), it stops transmitting; the surrounding neighbors detect this event and remove this node from their CRL lists. If the leaving node is a cluster head, rule (R1) will be activated and CMs re-execute this rule to elect a new CH and accomplish the necessary updates. Otherwise (case where the leaving node is not a CH), rule (R3) will be activated and CMs having it as forwarder (to access the CH in multi-hop fashion) will chose other paths to relay their information. The cluster leaving process is depicted in figure 4.6.

4.3.4.2/ Cluster Joining

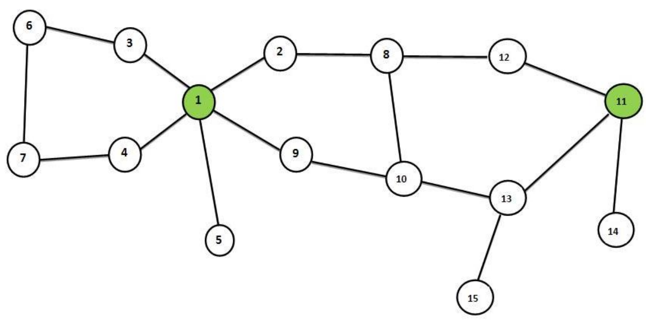
When a node i choose to join a cluster, it executes the initialization phase and broadcasts a join request to the nearby neighbors. These neighbors reply by sending an LSV_i beacon that includes the clustering information. The new node receives and stores it in its CLR_i list and computes its weight value W_i. If it has the highest weight among all the CLR_i members, then it elects itself as the new CH and broadcasts an announcement message, so that the neighbors can achieve the necessary changes (set the new node as their new CH). Otherwise, the new node joins the cluster of the weightiest CH in its k-hop neighborhood. Next, i selects the nearest node to the CH (node with the lowest RSSI value) as its new parent to relay the gathered data.



(**a**) Initial state

ID	W	Mych	WMych	Dist	|N(ID)|	Parent
1	3.2	—	—	0	5	—
2	-2	—	—	0	2	—
3	-1.5	—	—	0	2	—
4	-1.5	—	—	0	2	—
5	-4	—	—	0	1	—
6	0	—	—	0	2	—
7	0	—	—	0	2	—
8	0.7	—	—	0	3	—
9	-2	—	—	0	2	—
10	0.4	—	—	0	3	—
11	1	—	—	0	3	—
12	-1	—	—	0	2	—
13	0.7	—	—	0	3	—
14	-2	—	—	0	1	—
15	-2	—	—	0	1	—

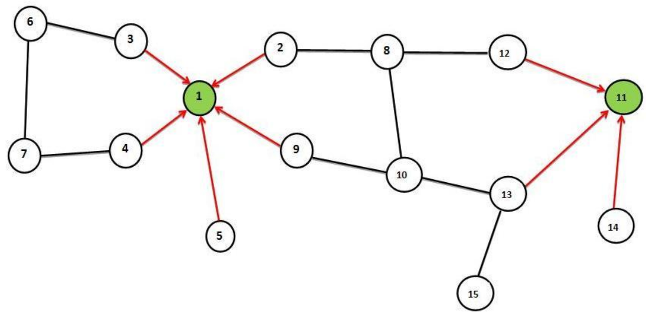
(**b**)



(**c**) CHs election

ID	W	Mych	WMych	Dist	|N(ID)|	Parent
1	3.2	1	3.2	0	5	//
2	-2	—	—	0	2	—
3	-1.5	—	—	0	2	—
4	-1.5	—	—	0	2	—
5	-4	—	—	0	1	—
6	0	—	—	0	2	—
7	0	—	—	0	2	—
8	0.7	—	—	0	3	—
9	-2	—	—	0	2	—
10	0.4	—	—	0	3	—
11	1	1	1	0	3	//
12	-1	—	—	0	2	—
13	0.7	—	—	0	3	—
14	-2	—	—	0	1	—
15	-2	—	—	0	1	—

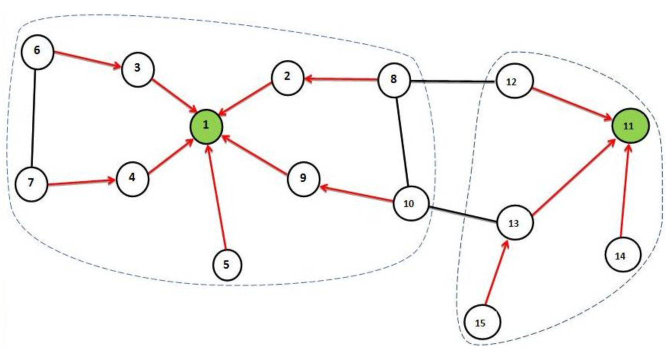
(**d**)



(**e**) Routing paths formation

ID	W	Mych	WMych	Dist	|N(ID)|	Parent
1	3.2	1	3.2	0	5	//
2	-2	1	3.2	1	2	1
3	-1.5	1	3.2	1	2	1
4	-1.5	1	3.2	1	2	1
5	-4	1	3.2	1	1	1
6	0	—	—	0	2	—
7	0	—	—	0	2	—
8	0.7	—	—	0	3	—
9	-2	1	3.2	1	2	1
10	0.4	—	—	0	3	—
11	1	1	1	0	3	//
12	-1	11	1	1	2	11
13	0.7	11	1	1	3	11
14	-2	11	1	1	1	11
15	-2	—	—	0	1	—

(**f**)



(**g**) Clusters formation

ID	W	Mych	WMych	Dist	|N(ID)|	Parent
1	3.2	1	3.2	0	5	//
2	-2	1	3.2	1	2	1
3	-1.5	1	3.2	1	2	1
4	-1.5	1	3.2	1	2	1
5	-4	1	3.2	1	1	1
6	0	1	3.2	2	2	1
7	0	1	3.2	2	2	1
8	0.7	1	3.2	2	3	1
9	-2	1	3.2	1	2	1
10	0.4	1	3.2	0	3	1
11	1	1	1	0	3	//
12	-1	11	1	1	2	11
13	0.7	11	1	1	3	11
14	-2	11	1	1	1	11
15	-2	11	1	2	1	11

(**h**)

Figure 4.5: Execution scenario of DC2HC clustering process.

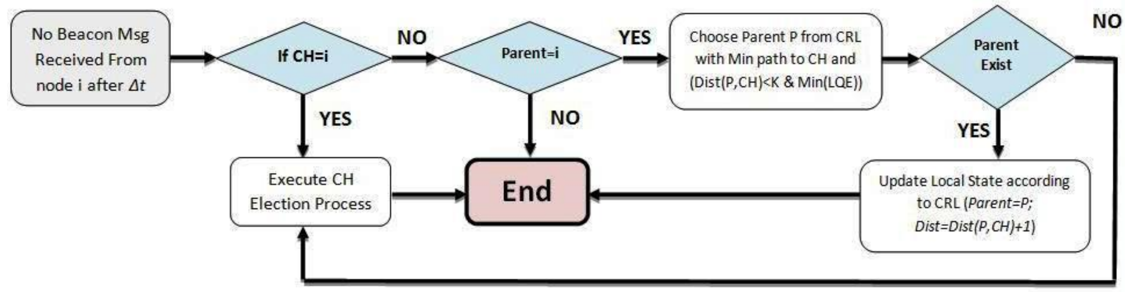


Figure 4.6: The cluster head leaving process.

In the case where no reply packet is received, i.e. the new node is within an isolated region, the new node i elects itself as a cluster head (to prevent isolated node scenario), updates its local state ($Mych_i = ID_i,\ Dist(i, Mych_i) = 0$) and broadcasts a $CH\_Announcement$ packet, since other nodes may have joined this region while i was performing the joining procedure. The cluster joining process is shown in figure 4.7.

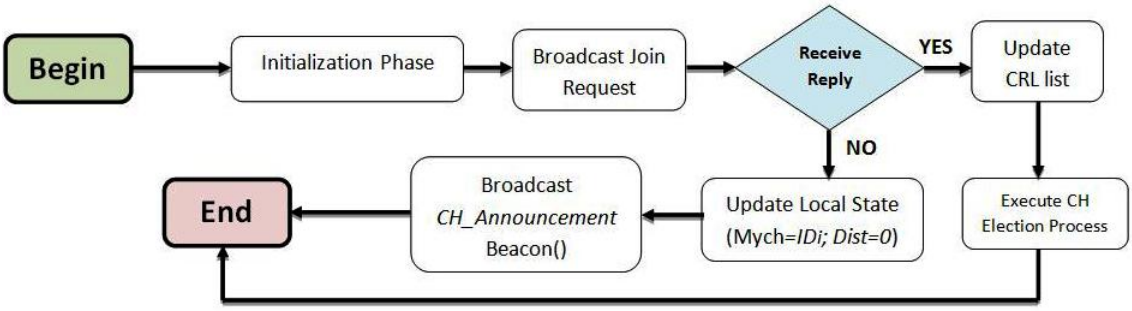


Figure 4.7: The cluster head joining process.

4.4/ TRANSMISSION RELIABILITY

Transmission reliability is a crucial criterion to improve the QoS of an application. The well-known available Link Quality Estimators (LQEs) are shown in figure 4.8. Among all these estimators, Radio Signal Strength Indicator (RSSI), Link Quality Indicator (LQI) and Packets Received Rate (PRR) are the most typical metrics used to estimate transmission quality [32]. RSSI measures the power signal of the received packets. LQI indicates the quality of the received packets based on the first eight bits of that packet. LQI is an efficient estimator, but the RSSI exhibits good results with a small number of measurements and converge faster than LQI. PRR is reviewed as the best indicator for link quality [145]. However, it requires more time and energy to perform a precise quality value. For these reasons, RSSI [146] is commonly chosen as the link quality estimator in this study. Lower RSSI value implies a weaker signal. It is quantified in decibel, the closer this value is to zero the better the signal is. For example, −50 dBm is a good signal and −75 dBm is fairly reasonable.

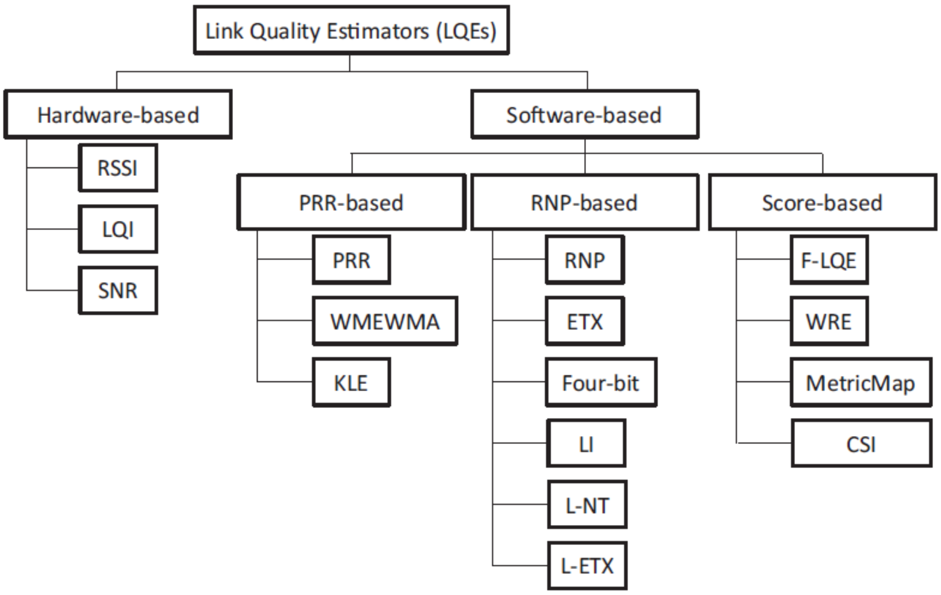


Figure 4.8: Available link quality estimators (LQE's) [14].

4.5/ Convergence proof

Since the network structure is dynamic, the clustering algorithms must have a quick convergence time [1]. The convergence time of the algorithm is evaluated in terms of rounds. Each round is quantified by a number of movements. In this section, we will prove that our approach requires a finite number of movements in the clustering process, which implies its convergence. We assume that the network starts from an arbitrary configuration.

Lemma [1]. *Let $i \in V$ be the node that has the highest weight value in G (i.e., $\forall j \in V - \{i\}$: $W_i > W_j \vee (\exists\ i \neq j : W_i = W_j \wedge ID_i > ID_j)$. i makes at most 2 movements (executes (R1) and (R3)).*

Proof. we show that node i executes a finite number of movements. □

We assume that the value of $Mych_i$ is not updated. Since i is the weightiest node in G, it has the highest weight among all its neighbors ($\forall j \in N(i) : W_i > W_j$). The rule $R1$ is enabled at node i, it will update the $Mych_i$ value (CH). Next, $R2$ is executed, i updates the distance $Dist(i, Mych_i) = 0$, afterward, no more rules are executed by node i, so it will

[1]Convergence time: represents the duration from when a node starts the execution of the clustering algorithm until the cluster is finally constructed. At the convergence point (stabilization point), no more operations are executed by the node until the end of the clustering process

not make another move. $Mych_i$ and $Dist(i, Mych_i)$ will have the same values during the setup phase.

Lemma 2. *Neighbors at distance $< k$ from the node with the highest weight $i \in G$ will execute a finite number of movements.*

Proof. We show that the k-hop neighbors of node i execute a finite number of movements (at most $2 \times k$ movements) before reaching the final state (Final state (or stable state): represents the state where the node has reached the convergence point. From this state, the node has a correct clustering variable (LSV) and all the algorithm rules are disabled for this node). □

After that, node i (the node with the highest weight) reaches a stable state (Lemma 1). Rule R1 and R3 will be enabled for its direct neighborhood $j \in N(i)$, these neighbors will update their $Mych$ value ($Mych = ID_i$), then execute R3 to update the distance $Dist(i, j)+ = 1$ (2 movements). After that, no more rule will be enabled in the neighborhood of i ($N(i)$ has reached a final state).

This execution will be repeated by all nodes $\in N_2(i)$. Each node $j \in N_2(i)$ will execute at most 2 movements ($R1$ and $R3$) then finish their execution. Hence, the same scenario is repeated for the rest of nodes until it covers all the nodes in the sub-graph $G' = \{i \bigcup N_{\leq k}(i)\}$ that contains i and its k-hop neighborhood. Thus, the k-hop neighbors of the node with the highest weight reach the final state after at most $2 \times k$ movements and will not be able to execute any more movements.

Lemma 3. *The stability of DC2HC is ensured.*

Proof. Since node i with the highest weight in the network executes a finite number of movements (Lemma 1) and the k-hop neighbors of node i also execute a finite number of movements (Lemma 2). This implies that the sub-graph composed by the set of nodes $\{i\} \bigcup N_k(i)$ reaches a stable state after a finite number of movements. Let $G'' = \{G \ / \ (\{i\} \bigcup N_k(i))\}$ be the graph obtained by removing the first stabilized sub-graph from G. The execution given above in G can be repeated in G'', the second sub-graph that contains the node with the highest weight in G'' stabilizes with the same reasoning, so all the nodes of the graph will follow this reasoning. This implies that the total number of moves executed by DC2HC is finite. □

4.6/ Time and space complexity

In the previous section, we demonstrated that the proposed approach terminates within a finite number of movements ($2 \times (k + 1)$ $moves$). However, the convergence complexity is not provided. In this section, the complexity of the proposed algorithm is examined in both theoretical and simulation-based analyses.

4.6.1/ Theoretical analysis

In the following, we assume the worst configuration from which our algorithm can start. Then, we study the complexity of DC2HC.

There is a known configuration from which DC2HC fails to provide good results (the worst-case scenario). This scenario is when nodes are related and organized in a straight line and their IDs are monotonically increasing or decreasing as illustrated in figure 4.9. In this configuration (network with low density), the maximum degree of each node $i \in G$ is: $0 < Deg(i) \leq 2$. In this case, network nodes will have the same weight and the election of cluster head depends on nodes IDs. Although, this configuration is unlikely to occur in a real-world network. It will allow the computation of the complexity in the worst case.

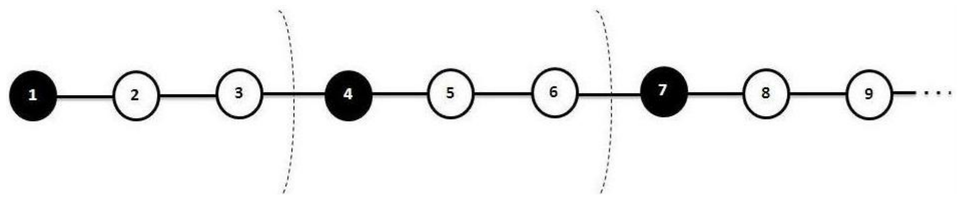


Figure 4.9: Worst-case scenario ($k = 2$).

In the following, we prove that the time complexity of DC2HC is at most $O(|V| + k)$ rounds, where $|V|$ is the number of nodes in the network.

Lemma 4. *Let i be the node with highest weight in G then:*

(a) After 2 rounds and in all following rounds, i is a cluster head ($Mych_i = ID_i$ and $Dist(i, Mych_i) = 0$).

(b) After $2 \times k + 2$ rounds and in all following rounds, the neighbors of node i at distance $\leq k$ form a cluster, where $\forall j \in N_{\leq k}(i) : (Mych_j = i) \wedge (Dist(i, j) \leq k)$.

Proof. First, we demonstrate that (a) is true. In the first round, as i is the node with the highest weight in G, it has the high weight value among all its neighbors.

$$\forall j \in V - \{i\},\ z \in N(i) : (W_i > W_j) \Rightarrow (W_i > W_z)$$

□

This node executes rule R1 to become a CH. In the second round, i executes rule R2 to update the distance value $Dist(i, Mych_i) = 0$. After that, no more rules are activated at node i, so it will not execute another move in the following rounds. Property (b) means that the sub-graph $G' = \{i \bigcup N_{\leq k}(i)\}$ that contains node i and its neighbors at distance $\leq k$ reaches a stable state after at most $2 + 2 \times k$ rounds. In the 3rd round, when node i finishes its execution, rule R1 will be enabled for its neighborhood. These neighbors join the cluster of node i. Next, they execute R3 to update their distance $Dist(i, Mych_i) = 1$ (round 4). After round 4, no more rules can be activated in the neighborhood of i, so they will not execute other movements in the following rounds.

The argument given above is repeated for the neighbors at distance $\leq k$ from i (i.e., neighbors of i will reach a stable state after $2 \times k$ rounds). Therefore, to form a cluster, two rounds are required to elect the CH and $2 \times k$ rounds to form the cluster. By induction, each node with the highest weight and its neighbors finish their execution after at most $2 + 2 \times k$ rounds.

Lemma 5. *The maximum number of clusters that can be generated with DC2HC in the worst case configuration is* $1 + n/(k + 1)$ *where* $n = |V|$.

Proof. DC2HC divides the graph G into several spanning trees where each tree has the node with the highest weight as its root (the cluster head) and its neighbors at a distance $\leq k$ represent the rest of this tree (cluster). Therefore, the minimum number of nodes that can be in a cluster is $k + 1$. In the worst-case scenario (the graph shown in figure 4.9), a related graph can contain $n/(k + 1)$ clusters. As each cluster is represented by only one CH, the graph can contain at most $n/(k + 1)$ cluster heads. One is added to the previous threshold to cover the case where $n \bmod (k + 1) > 0$. Thus, we obtain $[\frac{n}{k+1}] + 1$, which represents the maximum number of CHs formed by DC2HC in the worst-case configuration. □

Lemma 6. *DC2HC algorithm converges after at most* $2 \times (n + k + 1)$ *rounds.*

Proof. According to Lemma 4, the subset that contains the node with the highest weight $i \in G$ and its neighbors at a distance $\leq k$ $(i \bigcup N_{\leq k}(i))$ stabilizes after at most $2 \times k + 2$ successive rounds. The same process is repeated in the graph $G' = \{G \ / \ i \bigcup N_{\leq k}(i)\}$. As G contains at most $[\frac{n}{k+1}] + 1$ clusters (Lemma 5) and each cluster takes $2 \times k + 2$ rounds to reach a stable state. Thus, the proposed protocol requires at most $([\frac{n}{k+1}] + 1) \times (2 \times k + 2) = ([\frac{n}{k+1}] + 1) \times 2(k + 1) = 2(n + k + 1)$ rounds to form all the clusters in the network. □

Lemma 7. *DC2HC algorithm has a linear time complexity of* $O(n)$ *rounds.*

Proof. Based on Lemma 6, the upper bound of the time needed for the execution of DC2HC is $2 \times (n + k + 1)$ rounds. Thus, it is obvious that the time complexity (convergence time) of the algorithm is $O(|E| + k)$. As $|E| \geq k$, the proposed algorithm converges after at most $O(n)$ rounds. □

Lemma 8. *DC2HC algorithm has a linear space complexity.*

Proof. Each node has to maintain the 2-hop neighbors information in its CRL data structure. Therefore, the space complexity of each node will not increase as long as the local density remains constant. The space complexity of DH2HC is $O(d_2)$, where d_2 represents the 2-hop local density. As $d_2 \leq n$, the complexity of the proposed algorithm is linear, which is needed in a resource constrained environment. □

The number of nodes inside clusters depends upon the network density and the k-hop constraints. Indeed, in the case of complete graph, all nodes can be grouped within a singleton cluster. Hence, the maximum number of nodes in a cluster is n, where n is the number of nodes in the network. However, this situation may not often occur in real world networks. Therefore, by considering a related graph, the maximum number of clusters that can be generated by the proposal is $n/(k+1)$ (Lemma 5), thus, the minimum limit of members inside each cluster is equal to $\frac{n}{n/(k+1)} = \frac{n(k+1)}{n} = k+1$.

4.6.2/ Clustering Property

To prove that the proposed approach works properly. Clustering properties (safety and liveness) need to be satisfied.

4.6.2.1/ Safety property

The safety property ensures that network nodes are grouped into clusters and each cluster has only one CH to avoid overlapping between clusters.

Lemma 9. *The safety property is verified.*

Proof. Each ordinary node elects the node with the highest weight among its multi-hop neighbors as its cluster head and the variable $Mych$ holds only one value based on CRL list. Thus, a node can only belong to one cluster at a time and is covered by a unique CH. As a result, the safety property is satisfied. □

4.6.2.2/ Liveness Property

The liveness property ensures that the clustering progresses normally and reaches a final state after a finite time.

Lemma 10. *The liveness property is satisfied.*

Proof. According to Lemma 3 and Lemma 6, the proposed algorithm executes a finite number of movements and converges after at most $n + k$ rounds. Hence, the liveness property is satisfied.

4.7/ SIMULATION

4.7.1/ EXPERIMENTAL SETTINGS

In this experiment, we adopted the network configuration described in section 3.5. We add a new term used in this experiment. The network topology is composed of a variable number of nodes ($\delta \in [40, 1300]$) distributed across a square area of size Δ^2= 1000 $\times$ 1000 m^2, as shown in figure 4.10 (blue nodes represent the CHs and black nodes are the CMs). The BS is located in the top left corner of the network. Random distribution is assumed to generate a random network topology, typically used in clustering approaches in the literature to approximate a real deployment scenario. We assume that network nodes are quasi stationary, however, the network topology is dynamic owing to the reconfiguration events (cluster head election and re-affiliations) that may occur due to eventual node disconnection factors, such as: device breakdown, battery draining, lack of coverage, etc. The proposed scheme uses a self-stabilizing technique which is suitable for such dynamic scenario where disconnection events frequently occur. Indeed, the re-clustering process is locally driven and does not affect the entire network when disconnection events occur. Conventional sensor network usually uses wireless communication standard with low power consumption [147], such as IEEE 802.15.4 Zigbee (with a maximum transmitting distance of 100 m) or the 802.11n [45] (with a transmitting distance of 70 m). Therefore, based on the wireless communication standards used by regular wireless networks, in this experiment, we assume that network nodes have a transmission range of $Tr = 70$ m.

In this contribution, the CH rotation occurs when the weight of the current CH is exceeded by one of its k-hop neighbors, which means that the current cluster has undergone a considerable topology change. This mechanism avoids the quick alternation of CHs and further reduces the number of clustering messages generated. The experiment parameters are listed in Table 4.2. Increasing the number of communication hops inside each cluster generates long routing paths which may extensively increase the data transmission delay. Thereby, the performance of our clustering algorithm is analyzed by considering $k \in \{1, 2, 3\}$. The k-hop constraint ought to be specified depending on the application requirements to improve the network performance. The proposed scheme is experimented with varying maximum hop values and nodes density to analyze the network performance under different scenarios.

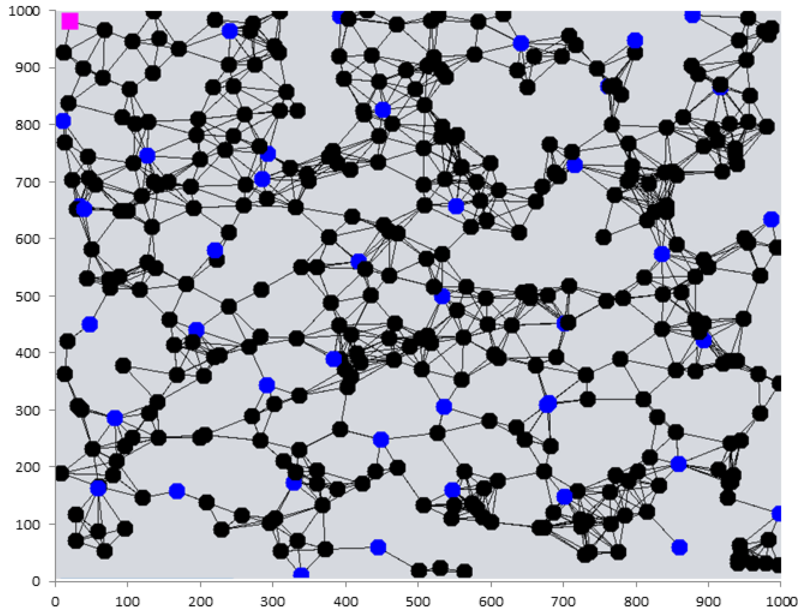


Figure 4.10: Network topology with 400 nodes deployed over a $1000 \times 1000\ m$ network

Table 4.2: Simulation parameters used in DC2HC setting (Using JUNG)

.

Parameter	Value
Network size (Δ^2)	1000 × 1000 m^2
Node density	$\delta \in [40, 1300]$
BS position	(20, 20)
Distribution of nodes	Random
connectivity model	Unit Disk Graph (UDG [131])
Transmitting range (Tr)	70 m
Maximum hop constraint k	$\{1, 2, 3\}$
α, β, γ	1/3, 1/3, 1/3
E_{elec}	50 nJ/bit
ε_{FS}	10 pJ/ bit/ M^2
ε_{Mfs}	0.0013 pJ/ bit/ M^4
Data packet size	100 bytes
Initial energy	1 Joule

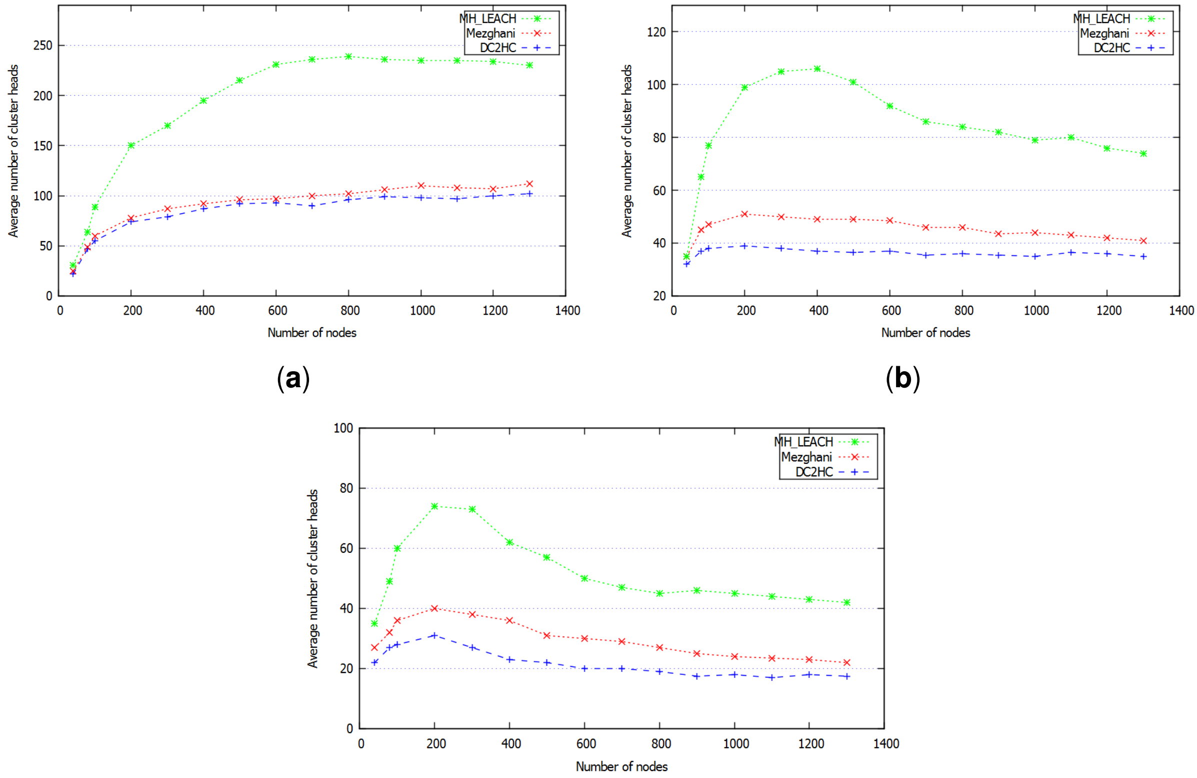
In the following section, the performance of our approach DC2HC is compared with two ubiquitous and recent protocols belonging to the same family of multi-hop clustering. MH-LEACH [89] and Mezghani protocol [116] are both used for the k-hop intra-clustering, and their primary objective is to reduce the waste of energy and the number of generated clusters. Due to these characteristics, these protocols are selected for the performance comparison to our proposed protocol.

Five parameters are used for the performance analysis of the simulated protocols, namely, CHs cardinality, energy consumption, number of exchanged messages, the average network lifetime, and the number of dead nodes. The same parameters are used to compare the three protocols. Each simulation result is the average of 10 measurements for each used metric with varying density and node distributions.

4.7.2/ Simulation Results

4.7.2.1/ Cluster Head Cardinality

The CH's cardinality stands for the average number of generated clusters. This metric enables the evaluation of clustering efficiency. Low CHs cardinality reveals good characteristics of the clustering scheme as it represents the number of communication channels established with the BS. Moreover, reducing the CHs cardinality restricts the usage of long-range communication channels, thus limiting the risk of congestion and reducing energy consumption. Figure 4.11 illustrates the average number of cluster heads produced by DC2HC, Mezghani [116] and MH-LEACH [89] protocols according to different nodes density $\delta \in [40, 1300]$. We observe that MH-LEACH generates the highest cardinality. Indeed, MH-LEACH uses a probabilistic technique for load balancing the CH task among nodes. This technique does not consider the residual energy or the surrounding environment of nodes. When the density converges from 200 to 800 ($\delta \in [200, 800]$) with $k \in \{2, 3\}$, the environment tends to be more connected, which enlarges the set of neighbors within the transmitting range Tr of each node. Therefore, in figure 4.11 (b) and (c), we observe that clusters cardinality of MH-LEACH start to decrease in the range between $\delta \in [400, 800]$ because the coverage of CHs regroups more nodes. Mezghani protocol considers the residual energy of nodes, which improves the cardinality when k = 1 (improved by an average of 43.7% compared to MH-LEACH). However, the clustering process only considers the average one-hop degree of nodes, accordingly, when $k \in \{2, 3\}$, the protocol slightly degrades in performance. The proposed scheme uses k-hop clustering, which minimizes the CHs cardinality and avoids having remote nodes within forced singleton clusters. According to figure 4.11, the proposed scheme exhibits better performance. It reduces the cardinality by an average of 18.3% and 62% when compared to Mezghani and MH-LEACH, respectively. Table 4.3 shows the average gain of DC2HC compared to Mezghani and MH-LEACH in terms of CHs cardinality.



(**a**) (**b**)

(**c**)

Figure 4.11: Average cluster head cardinality using the k-hop intra-clustering (**a**) $k = 1$, (**b**) $k = 2$, (**c**) $k = 3$.

Table 4.3: Synthesis of the average CH's cardinality gains compared with Mezghani and MH-LEACH

Clustering Algorithm	**Intra-cluster topology**		
	Single-hop	Two-hop	Three-hop
Mezghani	6.9%	21.1%	27%
MH-LEACH	55.3%	64.2%	67.3%

4.7.2.2/ Average Exchanged Messages and Consumed Energy

As wireless communication is the costliest operation in wireless networks and drains most of energy, the number of exchanged messages and the consumed energy are strongly related and has a considerable impact on devices lifespan. Figures 4.12 and 4.13, respectively, display the average energy exhausted by network nodes and the average number of exchanged messages during the clustering process according to a scaling density and considering $k \in \{1, 2, 3\}$. The curves' shapes show that the energy consumed by network devices is proportional to the number of exchanged messages. Usually, as the density increases, the energy consumed and the number of messages generated by the three approaches increase as well. The multi-hop communications mitigate the energy consumption by lowering the communication distance from CMs to their CH, especially when the size of the network scales. Figure 4.12 shows that whatever the number of deployed nodes, DC2HC consumes less energy than the other protocols. This improvement can be attributed to the small number of elected CHs and reduced CH rotation that decreases the number of clustering messages exchanged. The number of exchanged messages is reduced by 36.5% and 6.9% compared to Mezghani protocol and MH-LEACH protocol, respectively. The energy consumption is also reduced by an average of 21.7%, which is proper for increasing the network lifetime. Table 4.4 resumes the gains acquired by DC2HC when compared to Mezghani and MH-LEACH in terms of energy consumption.

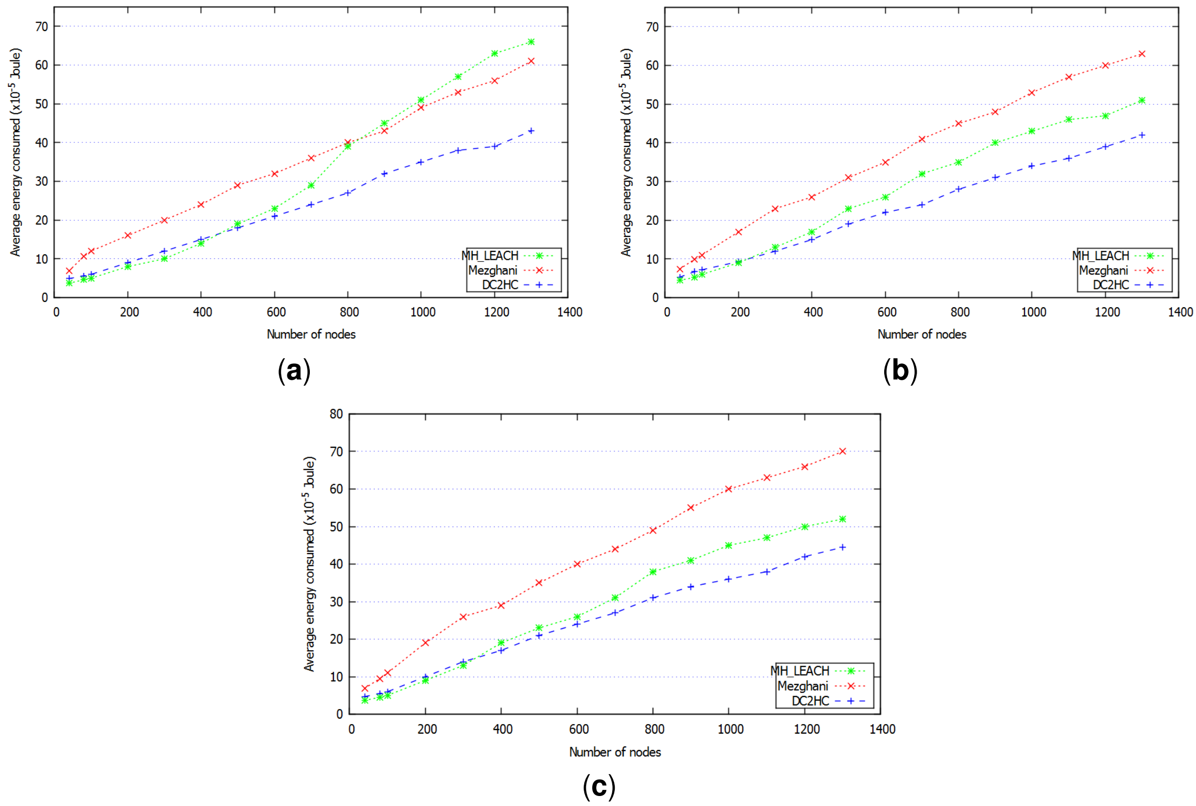


Figure 4.12: Average energy consumption of the three approaches with (**a**) $k = 1$, (**b**) $k = 2$, (**c**) $k = 3$.

4.7.2.3/ Average Network Lifetime

Several works define the network lifetime as the period from network initialization to the round when the last node dies [136; 89]. This parameter is important because it shows how long network nodes are able to execute the protocol. Moreover, the early death of

Table 4.4: Synthesis of energy consumption gains compared with Mezghani and MH-LEACH

Clustering Algorithm	**Intra-cluster topology**		
	Single-hop	Two-hop	Three-hop
Mezghani	28.7%	37%	37.1%
MH-LEACH	10.6%	13.8%	3.1%

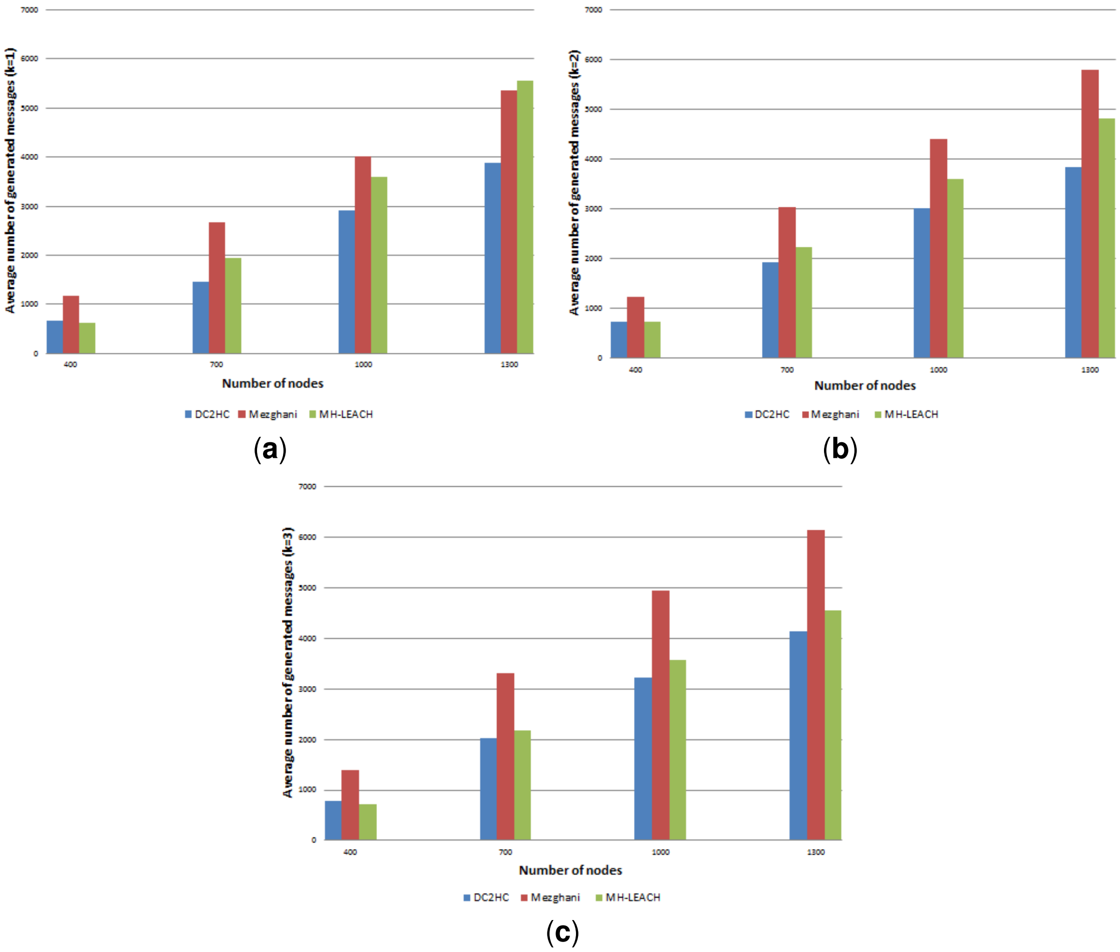
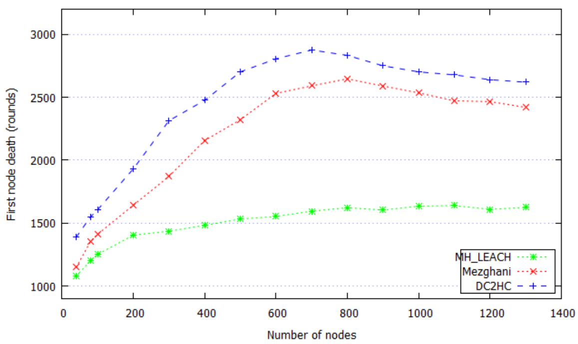


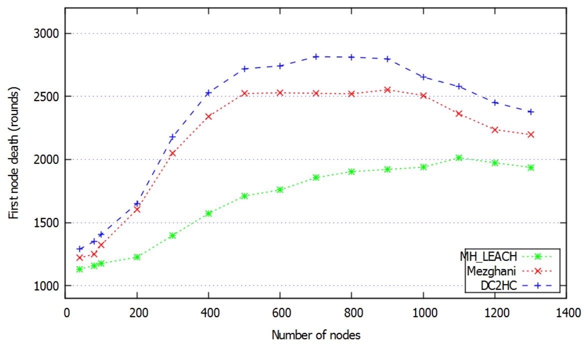
Figure 4.13: Average number of exchanged messages during the clustering process (**a**) $k = 1$, (**b**) $k = 2$, (**c**) $k = 3$.

nodes can lead to a disconnection of some parts of the network. Therefore, to measure the network lifetime, we considered the round at which the first node die (FND) and the last node die (LND). Figure 4.14 illustrates the time when the first node died according to different density levels. In MH-LEACH, nodes die rapidly because the protocol does not consider the residual energy of nodes. By considering Figure 4.14, when the density bypass 200 ($\delta \geq 200$) the FND round of MH-LEACH tends to stabilize between 1400 and

1600 rounds. Mezghani protocol reveals better performance compared to MH-LEACH (improved by 45.2%) due to the reduced number of generated CHs (Figure 4.11) and to the usage of Khalimsky theory [116] to elaborate an energy efficient cluster topology. According to figure 4.14 (a), DC2HC shows the highest durability, where the FND node reached 2876 rounds, whereas, in Mezghani and MH-LEACH, it only attained 2591 and 1565 rounds, respectively. Increasing the maximum hop constraint ($k > 1$) leads to the performance improvement of the three simulated protocols due to the use of multi-hop clustering, which reduces the energy consumption. The curve allure of DC2HC in figure 4.14 (b) shows a performance improvement compared to the other approaches. Indeed, the proposed scheme engenders a low cardinality of dense clusters based on the neighborhood TCR value. DC2HC applies the two hop neighbors' information to favor nodes located in a well-connected region to be CH. It uses a multi-hop routing tree within each cluster to improve the intra-cluster communications. Hence, the connectivity among clusters is improved, minimizing the energy consumed in transmission. Subsequently, D2MHC reaches a gain factor of 13.3% and 58.5% compared to Mezghani and MH-LEACH, respectively.



(**a**)



(**b**)

Figure 4.14: First node death according to different nodes density with (**a**) single-hop, (**b**) multi-hop.

Figure 4.15 shows the round at which all network nodes died versus the network density ($\delta \in [40, 1300]$) considering $k \in \{1, 2, 3\}$. In MH-LEACH, the role of CHs is rotated randomly and periodically, which increases the number of CHs; therefore, Mezghani protocol behaves better than MH-LEACH. However, using DC2HC allows a further extension of nodes' lifetime since it minimizes the set of cluster heads, thus reducing the waste of energy devoted to long range communications. Moreover, it limits the number of CH rotations, which reduces the clustering messages and the overall energy consumption offering the network a longer lifetime. Our proposed protocol outperforms Mezghani and MH-LEACH protocols in network lifetime by 26.3% and 13.3%, respectively.

Figure 4.16 depicts the total number of nodes that remain alive over execution time con-

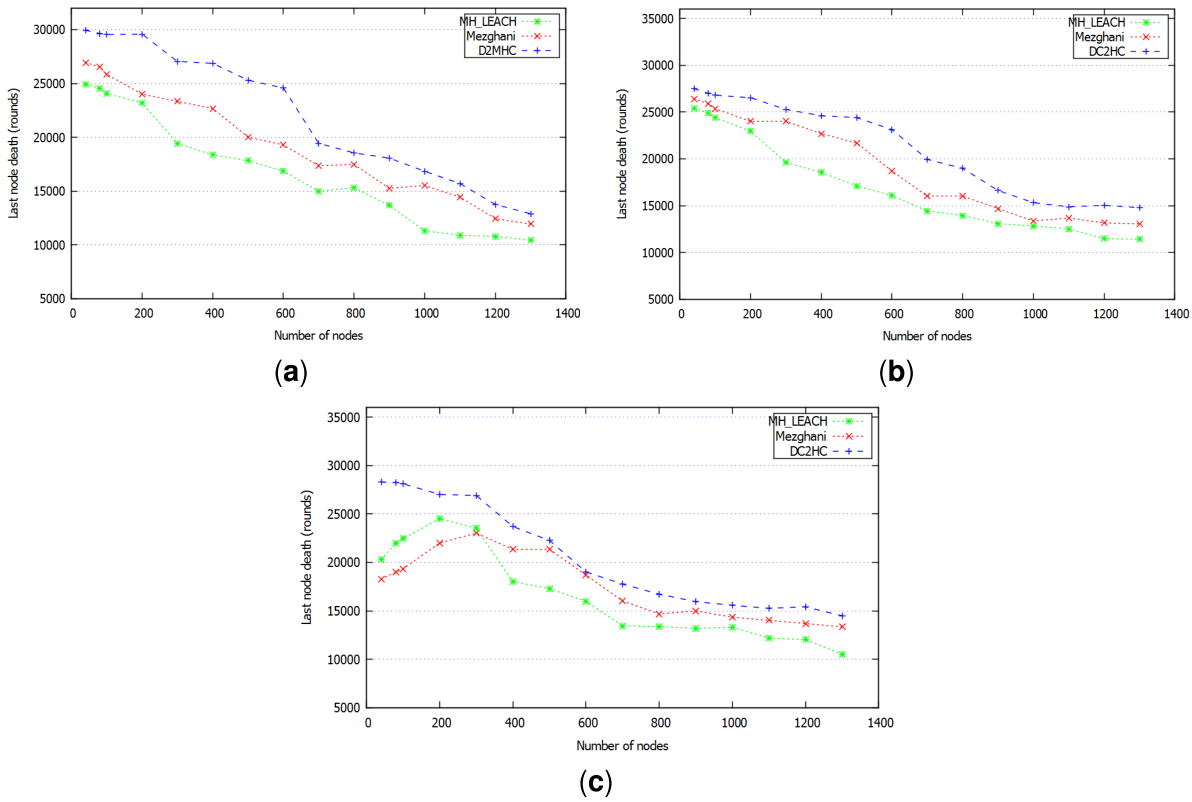
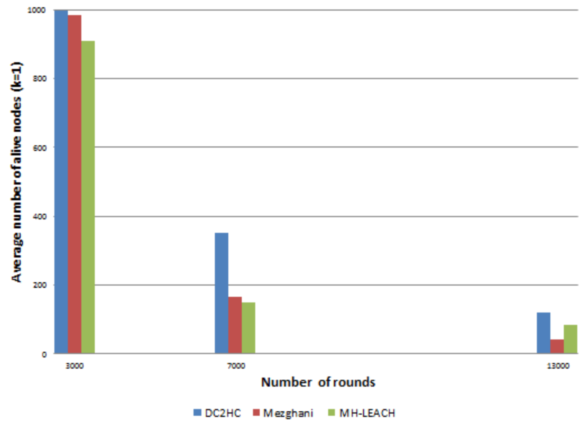


Figure 4.15: Last node death according to different nodes density with (**a**) $k = 1$, (**b**) $k = 2$, (**c**) $k = 3$.

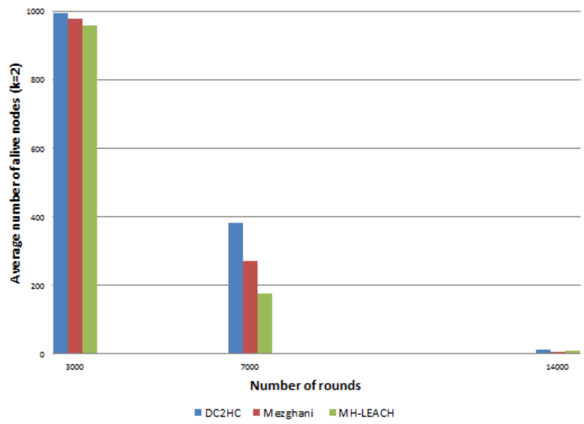
sidering the number of deployed nodes $\delta = 1000$. Although the performances of the three protocols initially look similar, we observe that DC2HC has more alive nodes than the other algorithm with the growth of rounds. MH-LEACH selects the set of CHs based on random probabilities and does not consider nodes' residual energy, which results in a faster death of network nodes. At the end of the clustering process, only a few nodes remain alive. Nevertheless, using our protocol allows network nodes to last for a longer period. Whereas with the other protocols, nodes died faster. Table 4.5 summarizes the gains obtained by DC2HC compared to Mezghani and MH-LEACH in terms of energy consumption.

Table 4.5: Network lifetime gains compared with Mezghani and MH-LEACH

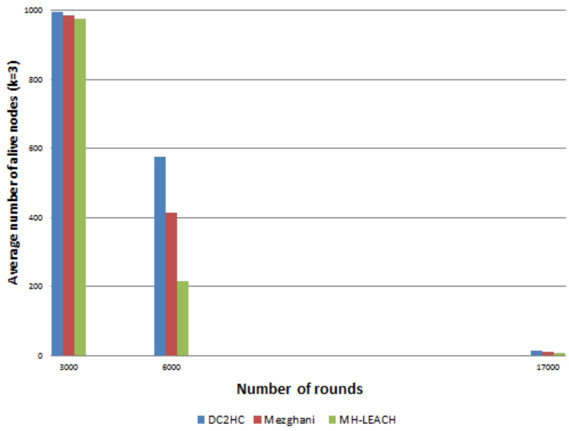
Clustering Algorithm	**Intra-cluster topology**			
	FND		**LND**	
	Single-hop	Multi-hop	Single-hop	Multi-hop
Mezghani	30.5%	24.2%	16.4%	10.2%
MH-LEACH	14.4%	14.7%	75.1%	41.8%



(**a**)



(**b**)



(**c**)

Figure 4.16: Average number of alive nodes over execution time with $\delta = 1000$ deployed nodes and (**a**) $k = 1$, (**b**) $k = 2$, (**c**) $k = 3$.

4.8/ Conclusion

Long distance communication channels used by cluster heads to reach the base station consume a significant amount of energy and accelerate the appearance of dead nodes and congestion problems. In this context, we presented a new distributed approach for k-hop intra-clustering called Distributed Clustering based 2-Hop Connectivity (DC2HC) for large networks. The proposal optimizes the set of cluster heads and extends the network lifetime. The cluster heads election is performed using the two-hop neighbors' topology knowledge to strengthen the cluster connectivity. It considers the residual energy and communications reliability to balance the energy consumption among larger clusters. We proved that the proposed approach has a linear time complexity of $O(|V|+k)$ rounds, where $|V|$ is the number of nodes in the network. Various simulation experiments have been performed to evaluate the performance of the proposed scheme with different parameters. The simulation results show the improved performance of DC2HC over the referred algorithms. Cluster heads cardinality is reduced by 40.1%, energy consumption is enhanced by 21.7% and network lifetime is extended by 20.8%.

5

DYNAMIC CLUSTERING TECHNIQUE USING AN ORIGINAL LONG-TERM VISION OF ENERGY OPTIMIZATION

5.1/ MOTIVATION

The energy storage system devoted to IoT devices can considerably restrict the network durability. Energy conservation is one of the prevalent and monolithic constraints in wireless networks. Commonly, devices' batteries represent the unique source of energy and therefore are central for the nodes' operation. The energy source problem undermines the integration of the traditional wireless network protocols into the IoT [95]. Therefore, the self-rechargeable battery concept has drawn much attention [95; 148]. This kind of batteries benefit from prolonging the lifetime of devices by fitting them with recharging systems [97](e.g. piezoelectric generators, RF harvesters, solar panels and wind turbines), which transform environment sources, such as body heat, foot hit, wind and sunbeams into electric energy [95]. However, the durability of rechargeable batteries is not unlimited and omitting the degradation aspect may lead to a high substitution and maintenance costs. Hence, preserving battery state of health (SoH) is crucial even with rechargeable batteries.

Although clustering techniques extend the network longevity, existing energy efficient works in this field, up to this point, mostly focus on short-term vision based on the state of charge (SoC) of non-rechargeable batteries and fail to address the impact of operational activities on the rechargeable battery deterioration in their clustering strategy. Such approaches do not optimize rechargeable batteries' lifespan (in the long-term). Indeed, the major parts of the IoT network devices use rechargeable batteries that age and degrade over time due to several factors (temperature, voltage, charging/discharging cycle, etc.). Therefore, it is essential to promptly detect these internal and environmental degradation factors to avoid network failures.

To meet the modern network requirement, saving energy must also be considered for rechargeable batteries to promote the development of ecosystems and reduces the material replacement frequency, which influences the green computing environment. Since battery aging is one of the prevalent failures that cause material replacement [149], considering the long-term exploitation of devices batteries in networking design can considerably decrease the carbon footprint and extend the network durability. Therefore, in this contribution, we discuss the lifetime extension using an original long-term energy optimization vision that enhances the network lifetime by taking into consideration the rechargeable batteries' degradation and their State of Health (SoH) during the clustering of the network. We describe a new Long Term Energy Optimization Clustering approach for IoT networks (LTEOC).

5.1.1/ Rechargeable and Non-Rechargeable battery lifespan

There is a often confusion when discussing battery lifetime because the lifespan for rechargeable and non-rechargeable batteries are defined in different manner. Non-rechargeable batteries, also called primary batteries [150] die and need to be altered after their initial charge is exhausted. Hence, the indicator for remaining battery life is the SoC, which represents a measure of the electrical energy amount stored and it indicates the charge that is left. This type of batteries cannot be recharged due to the irreversible electro-chemical reactions that take place inside the batteries. The Zinc–carbon batteries are one of the representative Non-rechargeable batteries [151]. In contrast, rechargeable batteries (also called secondary batteries [150]) can ensure multiple charge/recharge cycles (i.e. can be cyclically reused by recharging), enabling operation for extended intervals when combined with energy harvesting solutions like solar cells or thermal energy. The first commercial lithium-ion batteries (LIB, developed by Sony Corporation in 1991 [152]) trigger a revolution of the rechargeable battery market. Subsequently, the development of secondary batteries rises quickly, including the conception of nickel–metal hybrid batteries [153] and sodium-ion batteries [154]. Among these latter batteries, LIB exhibits the highest energy density and maintains an exceptional working performance, they are currently leading the rechargeable battery market and they are widely used in several areas [150].

In spite of their capability to be recharged, rechargeable batteries still have limited lifetimes and need replacement due to aging. Accordingly, instead of the SoC, we need to contemplate their SoH also, called Remaining Useful Life [7] (RUL), which is a figure of merit for physical battery condition. This parameter provides important information to forecast a battery malfunctioning and prevents network failures due to sudden nodes death.

Consequently, the SoH metric ought to be considered as it allows the implementation of energy aware applications in which devices operations can be altered depending on the battery state [97].

Fig. 5.1 outlines the difference between the short-term and long-term battery lifetime. The consideration of the SOC enables the improvement of the battery lifespan in the short-term (prolong the discharging cycle durability). However, in order to increase the number of recharging cycles and extend the battery durability in the long-term, the battery degradation and the SOH ought be regarded.

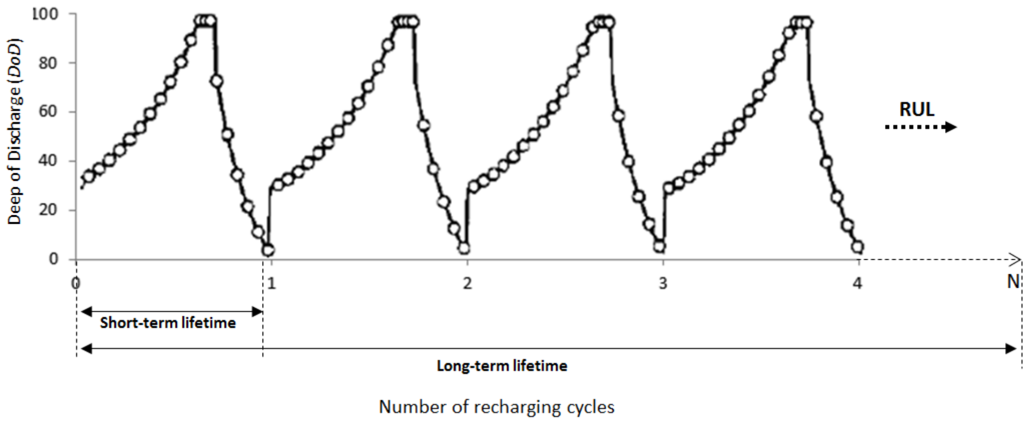


Figure 5.1: Illustration of the short-term and long-term rechargeable battery lifetime (N: battery death cycle)

5.1.2/ Battery degradation model

The SoH refers to the operational status of the battery and declines over time from 1 representing a healthy battery to 0 where the battery is completely inoperative due to degradation. Rechargeable batteries are impacted by diverse aging factors that lead to a considerable effect on their performance. These factors not only impact the performance of the battery, but they also reduce their lifetime [155]. Several aspects influence the battery life and provide the appearance of aging effects, such as calendar aging and cycle aging [156; 157].

Calendar aging stands for the battery's inherent degradation over time (in the long-term). It reflects the battery depletion caused by keeping the battery under given operating conditions, including the temperature and the charge degree [155]. These factors determine the aging speed over time. Meanwhile, the cycle aging is due to the charging/discharging cycles. It affects the degradation of the battery due to the complex composition and working process of the battery.

The selection of a proper model to estimate the current SoH is of primary importance.

Indeed, there are many methods proposed in the literature to determine the battery SoH [156; 157; 158; 9]. However, most of the proposed methods are too complex to run on low-cost micro-controllers. Therefore, in this work, we adopted the simplified battery degradation model used in [156].

In the adopted degradation model, the previous factors are used to estimate the battery additional degradation level during laps of time p. In this model, T_p defines the internal battery temperature during p, SoC_p represents the portion of available battery capacity, and DoD_p describes the portion of charge consumed during p. Hence, the State of Health loss during p, SOH_p is formulated by equation 5.1 as follows:

$$SOH_p = (S_p^{DoD} + S_p^t) \times S_p^T \times S_p^\sigma \tag{5.1}$$

Where $S_p^{DoD}, S_p^t, S_p^T, S_p^\sigma$ represent, respectively, the battery degradation contributions of DoD, calendar aging, battery internal temperature and the average SoC during the period p. These terms are computed using the following equations, where $|p|$ is the duration of the round.

$$S^{DoD} = \alpha_{DoD} \times DoD_p \times e^{(\beta_{DoD} \times DoD_p)} \tag{5.2}$$

$$S_p^t = \alpha_t \times p \tag{5.3}$$

$$S_p^T = e^{\left(\alpha_T \times (T_p - T_{ref}) \times \frac{T_{ref}}{T_p}\right)} \tag{5.4}$$

$$S_p^\sigma = \alpha_\sigma \times e^{SOC_p - \sigma_{ref}} \tag{5.5}$$

In our simulations, we considered the values of the model presented in table 5.1.

Table 5.1: Stress model parameters for the battery degradation

Parameter	**value**
α_{DoD}	0.05
β_{DoD}	0.03
α_σ	1.04
σ_{ref}	0.50
α_T	6.93E-2
T_{ref}	25°C
α_t	4.14E-10/s

Figure 5.2 illustrates an example of SOC variation using a rechargeable battery and evaluates the degradation level (SoH) according to the model of equation 5.1. At the beginning, the nodes SoH level is set to 100%.

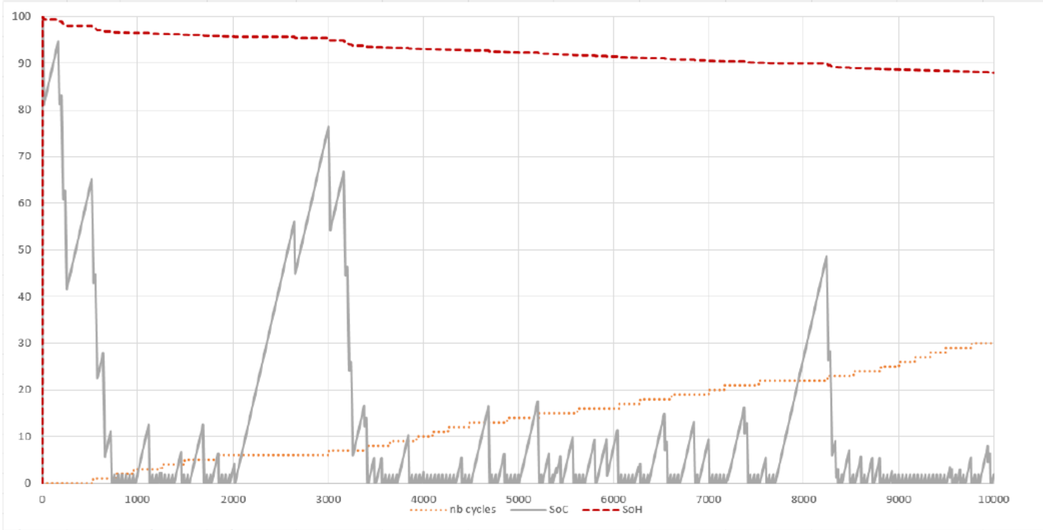


Figure 5.2: short-term vs. long-term optimization

5.2/ Contribution

In this part, we present a novel distributed multi-hop clustering technique that integrates an original long-term vision of energy optimization for IoT network by introducing both the battery State of Health (SoH) and the SoC in the clustering process. The proposed scheme effectively manages the network. It uses the devices' available energy resources in to optimize devices' battery lifetime. In addition, it decreases the number of BS connections required to get internet access.

In this contribution, we introduce a new dynamic clustering technique using an original long-term vision of energy optimization for IoT network (LTEOC: Long Term Energy Optimization Clustering), by considering the rechargeable battery aging aspect during the clustering process. In this solution a node uses its direct neighbor's information to perform multi-hop clustering. Each network node is assigned a weight value; the node with the highest weight is more potential to be elected as a CH. Neighbors at a distance of at most k-hop from the elected CH become Cluster Members (CMs). Nodes weight is performed by combining of three main metrics: the Adjacent Connectivity Ratio, battery SoC ratio and the rechargeable batteries SoH.

- **Adjacent Connectivity Ratio (ACR)**: this metric considers the relative connectivity proportion of a node with its neighborhood. A node with high connectivity ratio means that both the current node and its surrounding neighbors are well connected with many other nodes. Thereby, electing the current node as CH is appropriate to improve the network connectivity and reduces isolated nodes, as it will cover a large number of nodes. This metric is similar to the TCR metric used in the previous contribution (illustrated in Fig. 4.1 chapter 4), however, the ACR metric considers only the direct neighbor's connectivity to reduce the computation complexity. The ACR connectivity metric is calculated as in equation 5.6.

$$\Psi_i = \frac{\left|\left[\cup_{j\in|N(i)|} N(j)\ \right] \cup N(i)\right|}{|N(i)| + 1}$$

$$ACR_i = |N(i)| - \Psi_i \tag{5.6}$$

- **Battery State of Charge ratio (SoC)**: It represents the battery level of charge relative to its capacity. The ratio of the remaining energy of a device i is computed using equation 5.7:

$$SoC_i = \frac{E_{resid}}{E_{init}} \tag{5.7}$$

Where E_{init} is the initial charge of the current device battery and E_{resid} is the residual energy.

- **Rechargeable battery State of Health (SoH)**: the value of SoH for a battery varies over time due to degradation. It decreases from 1, representing a healthy battery (100%), to 0, referring to fully dead battery (0%). Stress factors affecting the battery aging also reduce the SoH since the metric is related to the battery deterioration. The SoH is described by the following equation:

Therefore, node's weight W_i is evaluated (based on the previous metric) as follow:

$$W_i = \alpha * ACR + \beta * SoC + \gamma * SoH \ \wedge \ \alpha + \beta + \gamma = 1 \tag{5.8}$$

The proposed solution is based on the following assumption: each node has a limited view of the network. It is aware of its local information and the information of its surrounding neighbors. Each node maintains a list of variables that represent its local state called LSV (Local State Variable). The LSV structure contains: the node identifier (ID), its weight W_i, $Mych_i$ (the relative CH of the current node), CH_weight (the weight of the relative CH), $Dist(i, Mych_i)$ that indicates the number of hops to the CH, $Deg(i)$. The node also has a clustering record list (CRL) in which it stores the $LSV's$ information received from its neighborhood. $LTEOC$ is composed of three main phases: the initialization phase, the cluster heads election phase, and the maintenance phase.

5.2.1/ Initialisation phase

.

A periodic $HELLO$ message is broadcasted by network nodes to identify the local neighborhood. Nodes that receive a $HELLO$ message, updates their local state and reply by sending an LSV beacon. After receiving an LSV beacon from neighbors, the concerned node updates its CRL list and recalculates its weight, then rebroadcasts an LSV message with the current information. With this scheme, nodes can preserve consistent information about their surrounding environment.

5.2.2/ Cluster heads election phase

The pseudo code for the clustering process is depicted in algorithm 4. In order to elect a cluster head, a node i checks for the largest weight value ($Highest\_w$) among its neighbors and the CHs that dominate its neighbors (rule R1) using the CRL list. If the weight of the current node W_i exceeds the $Highest\_w$, then i elects itself as the new cluster head (rule R2). Next, it broadcasts an announcement message with the new CH information. Otherwise, node i elects the node with the highest weight as its new CH. In case of a tie, the node with the highest ID is elected. The chosen CH must not be at a distance that exceeds k hops to respect the maximum hop constraint and avoid the generation of poorly structured clusters. If the chosen CH does not meet these conditions, this CH is excluded from the election. This process is repeated until a suitable CH is found. If no node in CRL meets these requirements, i self-elects as a CH to prevent an isolated node scenario. After the CH election, nodes use the local information received to designate the shortest path leading to their CH (rule R3).

Figure 5.3 illustrates the initialization phase and the clusters' formation process of a node that decides to integrate the network.

5.2.3/ Maintenance phase

The re-clustering process is locally driven and does not cripple the whole system when disconnection events occur. The frequent shared control information ensures a faster feedback about disconnection events. Therefore, the resulting structure is more efficient and stable. When a node leaves its cluster due to a disconnection (stops transmitting), the nearby neighbors discover this event and delete the outgoing node from their CRL lists.

Algorithm 4 Cluster Head election phase

Code for each node *i*

Variables:
$N(i)$: set of neighbors node of i
W_i: weight of the node i
$Highest\_w$: temporary variable indicating the node with the highest weight
Output
CRL_i: clustering record list of i
$Mych_i$: the relative CH of i
$Dist(i, j)$: the minimal distance between i and j

(R1): If $Highest\_w \neq Max(\{CRL[ID_j][W] \mid j \in (N(i) \vee Mych_{N(i)}) \wedge Dist(i, j) < k\} \bigcup \{W_i\})$ **then**

$Highest\_w = Max(\{CRL[ID_j][W] \mid j \in (N(i) \vee Mych_{N(i)}) \wedge Dist(i, j) < k\} \bigcup \{W_i\})$
$Mych_i = ID_j$

(R2): If $(Mych_i = ID_i) \wedge CRL[Mych_i][Dist] \neq 0$ **then**

$Dist(i, Mych_i) = 0$
$CRL[Mych_i][Dist] = 0$
Broadcast $CH\_announcement$ beacon

(R3): If $(Mych_i \neq ID_i) \wedge Dist(i, Mych_i) \neq Min(Dist(j, Mych_j) \mid j \in N(i) \wedge Mych_i = Mych_j) + Dist(i, j)$ **then**

$Dist(i, Mych_i) = Min(Dist(j, Mych_j) \mid j \in N(i) \wedge Mych_i = Mych_j) + Dist(i, j)$
$CRL[Mych_i][Dist] = Dist(i, Mych_i)$

Update ACR, SoC, SoH values using equations 5.6, 5.7, 5.1
Update $W_i = \alpha * ACR + \beta * SoC + \gamma * SoH$
Broadcast LSV_i Beacon

If the leaving node is a CH, CMs will execute rule (R1) to elect a new CH and perform the routes updates. Otherwise, rule (R3) will be executed to establish new routing path leading to the CH. On the other hand, when a node chooses to integrate the network, it broadcasts a join request to the proximate neighbors. Then, it executes the initialization and the cluster head election phases to integrate a cluster.

Figure 5.4 shows an execution example of LTEOC in a small network represented by 6 nodes (green nodes represent the CHs, node 6 joins the network). The blue tables show nodes current parameters.

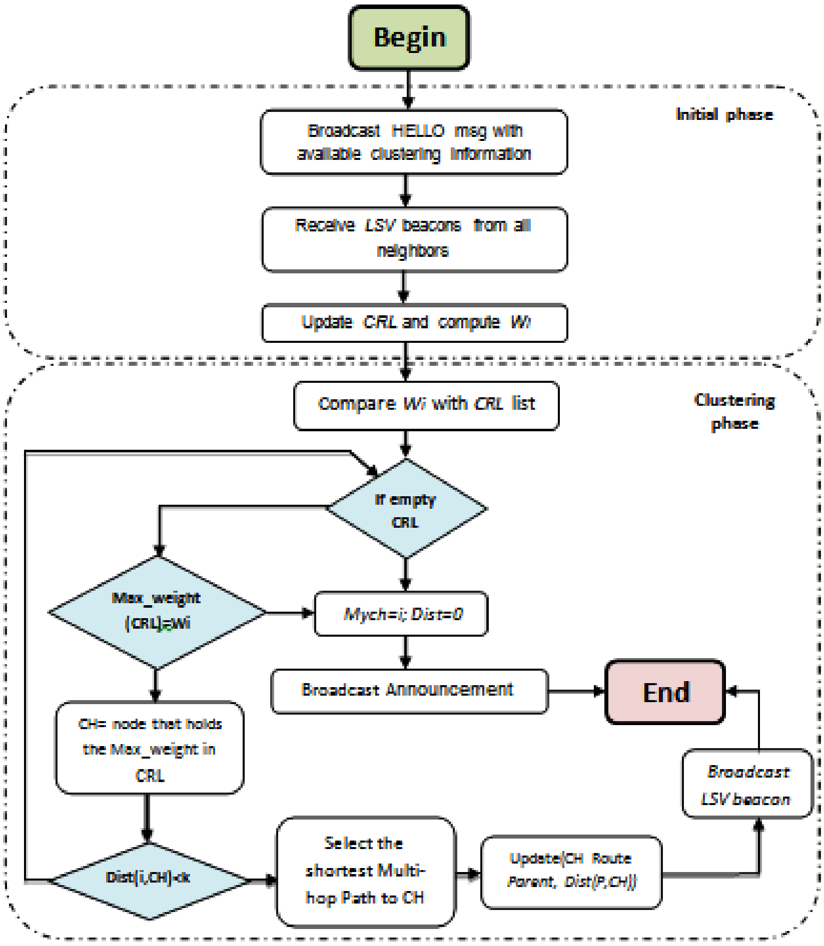


Figure 5.3: The initialization and the clusters formation process

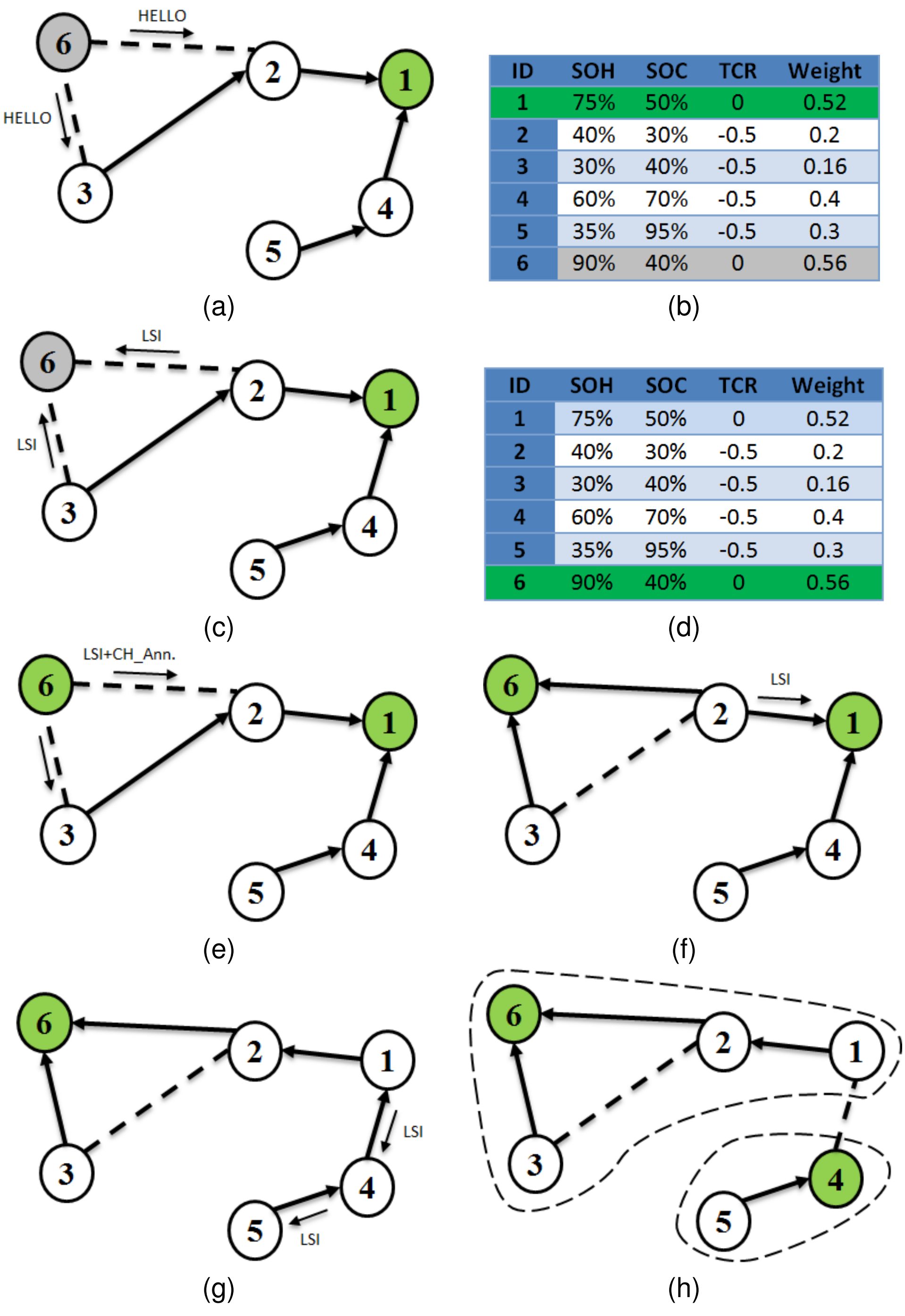
5.3/ Simulation

5.3.1/ Simulation parameters

In this experiment, we adopted the network parameters described in section 4.7.1. The experiment parameters are specified in Table 5.2. The power supply system is assumed to be outsourced, where devices are equipped with a mobile charger that responds to the charging requests. The temperature of network devices is also taken into consideration in the battery aging model and varies according to the task undertaken by the device. In each round, network nodes either send or receive packets from their environment.

We compare the performance of the scheme to that of DWEHC [159] and RinTRAR [119]. The two protocols belong to the same family of multi-hop clustering. The comparison is based on the following metrics:

- **Cluster head cardinality**: usually, smaller cardinality exhibits the efficiency of the



ID	SOH	SOC	TCR	Weight
1	75%	50%	0	0.52
2	40%	30%	-0.5	0.2
3	30%	40%	-0.5	0.16
4	60%	70%	-0.5	0.4
5	35%	95%	-0.5	0.3
6	90%	40%	0	0.56

ID	SOH	SOC	TCR	Weight
1	75%	50%	0	0.52
2	40%	30%	-0.5	0.2
3	30%	40%	-0.5	0.16
4	60%	70%	-0.5	0.4
5	35%	95%	-0.5	0.3
6	90%	40%	0	0.56

Figure 5.4: Execution scenario of LTEOC in a small network

proposed scheme, it constitutes the number of long-range and energy costly communication links established between the CHs and the BS. It also outlines the number of internet connections established.

- **Average battery life cycle**: depicts the average number of rounds required before the first network device drains out its energy and it needs to be recharged. This

Table 5.2: Simulation parameters used in the experiment setting

Parameter	**Value**
Network size (Δ^2)	1000m $\times$ 1000m
Neighborhood density	$\Phi \in [0, 30]$
Distribution of nodes	Uniform random
connectivity model	Unit Disk Graph (UDG [131])
Transmitting range (Tr)	30 m
BS position	(50, 50)
Maximum hop constraint k	$\{1, 2, 3\}$
E_{elec}	$50 nJ/bit$
ε_{FS}	$10\ pJ/bit/M^2$
ε_{Mfs}	$0.0013\ pJ/bit/M^4$
Data packet size	$50 bytes$
Initial SoH	1 (100%)
Initial energy	1 Joule

metric enables the estimation of the energy consumption performance of devices in the short-term (one battery cycle).

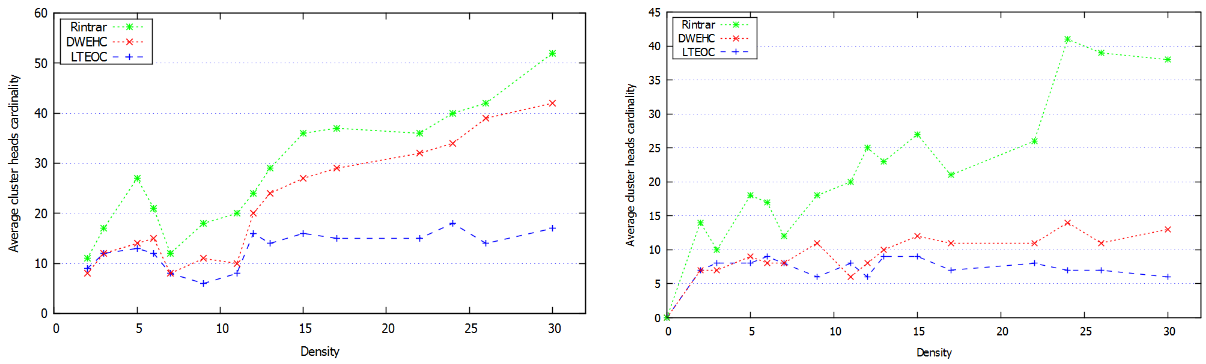
- **Average network lifetime**: Represents the average number of rounds needed before the rechargeable battery of the first network device is entirely dead and requires to be substituted. It allows the evaluation of the network duration and measures the battery degradation behavior in long-term.

- **Average number of dead nodes**: this metric is an important parameter to illustrates the longevity of the proposed technique and quantify the average required time before a rechargeable battery is totally dead and needs to be replaced. It develops a vision of the material replacement frequency.

5.3.2/ Simulation results

5.3.2.1/ Cluster Head Cardinality

Figure 5.5 depicts the average cardinality of network CHs generated by the three approaches. The connectivity metric (ACR) used to estimate the weight in the clustering procedure allowed the reduction of the average cardinality by 18% and 60.2% compared with DWEHC and RINtraR, respectively. The connectivity becomes difficult to guarantee when the network expands; therefore, multi-hop communications become essential. Figure 5.5 (b) shows that increasing the maximum hop constraint k lowers the cardinality; indeed, it permits more nodes to be covered within a single CH in a multi-hop mode. RINtraR generates the largest set of cluster heads due to the random probabilities exploited in the CHs election phase.



(a) (b)

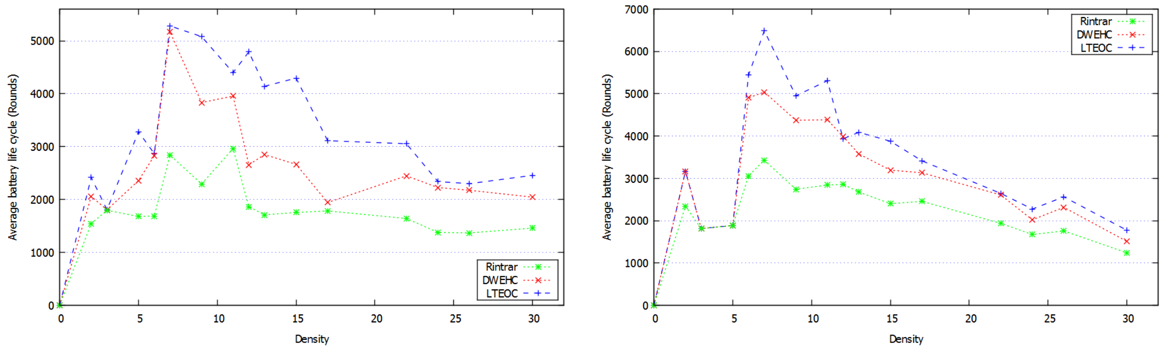
Figure 5.5: Average Cluster Head Cardinality: (a) k=1. (b) k=2.

5.3.2.2/ Average battery life cycle

Figure 5.6 illustrates the average battery cycle lifetime in term of rounds of the three algorithms. With low-density ($\Phi \leq 6.3$), the discharging of device energy is approximately equal. A curve spike is yielded when the connectivity increases ($\Phi = 7.5$). However, when the density increases ($\Phi \geq 12.5$), the devices' lifespan decreases. At this point, CHs attend to regroup more CMs and relay more data packets which induces a fast draining of residual energy. Implementing ACR and SoC metrics enables the devices using our approach to better conserve their energy. The battery cycle lifetime is improved by 16.9% and 58.2% in comparison to DWEHC and RINtraR, respectively. The use connectivity metric (ACR) in our approach allows the election of CHs with well-connected neighborhood, which intensifies the cluster coverage area and eases the CHs number. In DWEHC, the limited number of children per node leads to a boosted number of CHs in a dense network. LTEOC declines the usage of long-range communication to conserve the residual energy.

5.3.2.3/ Average network lifetime

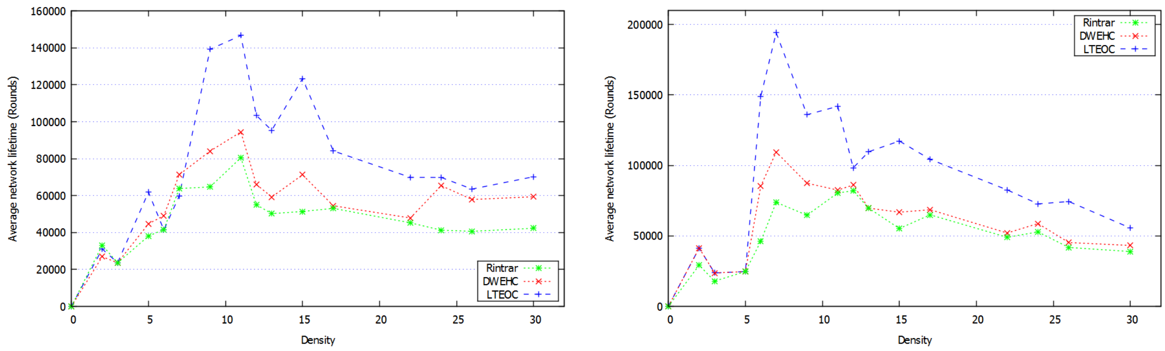
An analysis of the average network lifetime in long-term of our scheme versus DWEHC and RINtraR is depicted in figure 5.7. Initially, with low-density ($\Phi < 6$), many isolated nodes appear and are self-elect as CHs, in order to be capable of transmitting their data directly to the BS. Accordingly, at this stage, the results of the three protocols are quite similar. However, when $\Phi > 7.5$, nodes that aggregate information toward the BS will have more packets to rely; hence, their temperature rises rapidly, which leads to faster battery degradation and restrict the devices' lifetime. We notice that during simulation, the con-



(a) (b)

Figure 5.6: Average battery life cycle: (a) k=1. (b) k=2.

sideration of the batteries SoH in the clustering process manages the battery temperature and balances more efficiently the roles of CHs among network devices in long-term. Usually, when the density increases, many cliques emerges in the graph; using SoH metric, the CH role tends to change uniformly among the cliques' members, which balances the temperature reasonably among devices. With $k = \{1, 2\}$, the proposed approach allows network devices to endure for long periods of time and exhibit an average improvement of 37.6% and 51.5% compared to DWEHC and RINtraR. When $k = 3$, network clusters will contain more members in multi-hop fashion, and the load managed by CHs is balanced, which expands the battery aging. The proposed scheme boosts devices' lifetime by an average of 53.6% in comparison with DWEHC and RINtraR. However, with multi-hop communication, latency also increases, which is inappropriate for real time applications.



(a) (b)

Figure 5.7: Average network lifetime in long-term: (a) k≤2. (b) k=3.

5.3.2.4/ Average number of dead nodes

LTEOC incorporates battery degradation metric into the clustering process, allowing devices network devices to stand for long period. Indeed, during the simulation, we observe that considering battery SoH attends better balance function of the CH among network nodes and manages battery temperature between adjacent devices. Therefore, the rapidity of the battery aging process declines, and the endurance of network devices increases in the long-term. Figure 5.8 illustrates the average number of completely dead nodes during the simulation (quantified in rounds) by using our proposal versus DWEHC and RINtraR; the initial number of alive nodes is 240 ($\Phi = 30$). The usage of the residual energy in DWEHC permits devices using this protocol to last for longer periods as compared to RINtraR (an average improvement of 5.8%). The contribution produced the smallest set of dead nodes (17.4% and 23.6% lower that DWEHC and RINtraR), which reveals the impact on the energy optimization of our perspective in long-term.

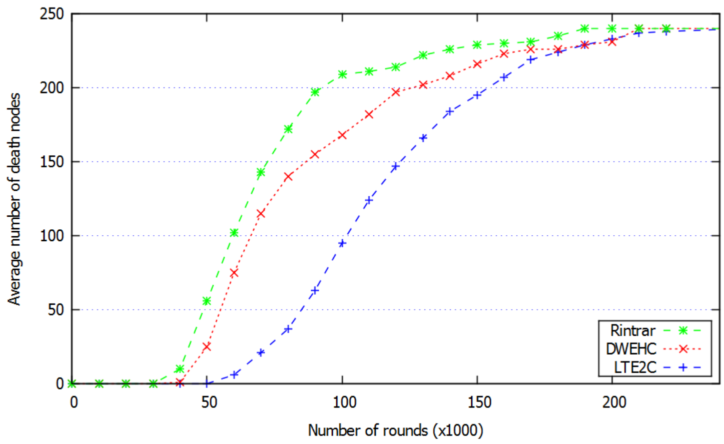


Figure 5.8: Average number of dead nodes according to the number of rounds (k=2)

5.4/ Conclusion

In this chapter, We treat how a long-term vision of energy consumption can be performed by taking into account the devices state of health and especially the rechargeable battery status. We demonstrate how the consideration of the battery aging characteristic can influence on the devices lifetime and on the network longevity in general. This regard can considerably reduce the materials substitution frequency and the carbon footprint in telecommunication systems. We presented a multi-hop clustering technique based on a long-term vision of energy optimization for IoT network (LTEOC) by considering the aging aspect of devices rechargeable battery to reduce the energy harvesting and enhance the network lifetime in long-term. The attained results illustrate that the aging based approach bypasses the results obtained with other methods regarding cluster heads cardinality, device battery life cycle and network durability. This work presents a first step in integrating the battery degradation aspect in the wireless networking protocols.

6

Battery State-of-Health Prediction based Clustering for Lifetime Optimization in IoT Networks

6.1/ Motivation

Conventional works either focus on static battery-powered networks (with one discharging cycle) or merely consider the current state of charge (SoC) of the batteries in order to improve the network lifetime in the short term. In this contribution, we propose a novel Long-term Energy optimization Clustering Approach based on battery State Of Health (SOH) prediction, called $LECA\_SOH$. The objective is to predict the impact of Cluster Heads election on the rechargeable batteries SoH before applying the clustering. $LECA\_SOH$ fosters the selection of the nodes, leading to the reduction of battery degradation during the future rounds, which in turn extends the system lifetime. In addition, considering the long-term exploitation of devices batteries in networking design can considerably decrease the carbon footprint.

Noteworthy that the work on rechargeable battery degradation modeling is an outcome of the electronics experts community and is not exploited enough by the networking community. Indeed, A great variety of research works in the literature have shown a particular concern to the modeling of the rechargeable battery degradation process [97; 155; 148; 160]. These researches deliver a promising concept of battery degradation for vehicular systems [7]. However, this concept was not introduced to the wireless networking systems. The main originality of this approach is to exploit the existing knowledge on the rechargeable battery degradation process in the conception of clustering and networking protocols for rechargeable wireless networks.

To illustrate the intended idea of $LECA\_SOH$ and clarify its novelty compared to the previous contribution $LTEOC$ (presented in the previous chapter 5), lets consider the routing example illustrated in Fig 6.1. In this routing scheme, four potential CH's can relay the data of the node S to the BS. With conventional energy aware approaches, either the node 1 is selected as CH because it residual energy is the highest ($SOC = 80\%$) or the node 2 because it consumes the lowest amount of energy ($\Delta_{SOC} = 15\%$). However, for a long-term energy efficient perspective, the optimization of batteries health leads to select node 3 or 4. Indeed, these nodes present a better SOH (node 3 with $SOH = 53\%$ and node 4 with $SOH = 51\%$). In this scenario, according to the previous contribution ($LTEOC$), node 3 will be selected because it presents a better SOH ($SOC = 53\%$). However, $LECA\_SOH$ grants node 4 as CH because this latter will endure less battery health degradation after the data aggregation ($\Delta_{SOH} = 0.4\%$) compared to the other nodes. Therefore, the predicted state of health of node 4 after routing is the best, which is more convenient for a long-term durability.

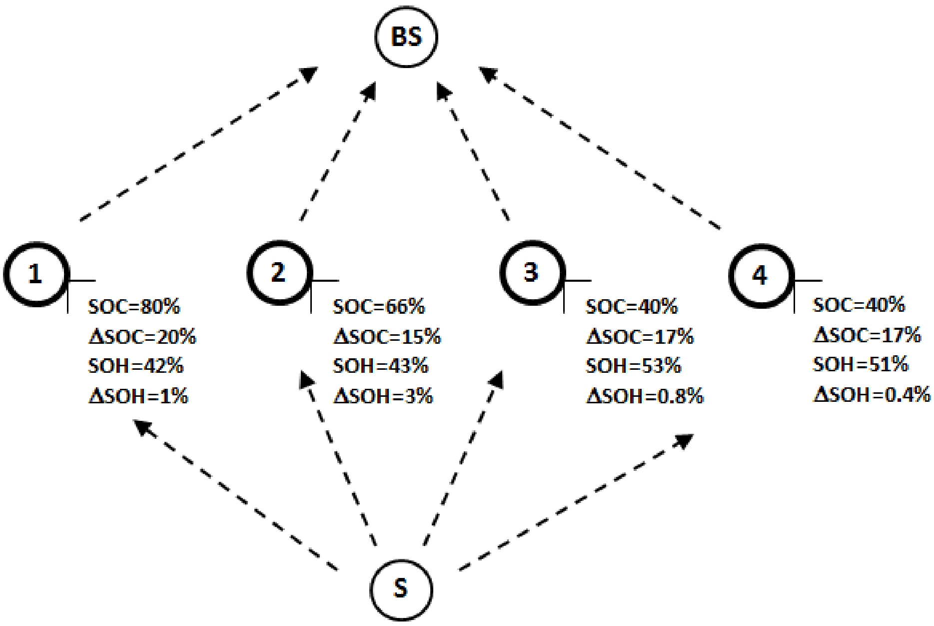


Figure 6.1: Short-term vs long-term energy efficient routing scheme

6.2/ Battery Internal Temperature Evaluation

In this contribution, we used the network model and battery degradation model described in section 4.7.1 and 5.1.2 respectively. In this section, we detail the battery internal temperature evaluation model used in this work.

Increasing the battery operating temperature above the recommended scope intensifies the aging process and leads to a fast battery degradation. Typically, the acceptable temperature scope of a Lithium-Ion Battery (LIB) is between [-20°C - 60°C] [150]. When the temperature is out of these interval, the battery degrades faster with a high risk of causing a safety issues including fire and explosion [150]. The acceptable temperature range should lie within [20°C - 40°C] to ensure a proper balance between performance, battery life and safety [161]. Low temperature affects the properties of electrolyte in LIB's. Indeed, with the diminution of temperature, the viscosity of the electrolyte increases, which reduces the ionic conductivity [150]. Consequently, the internal resistance will rises owing to the increase in the impedance of the directional migration of chemical ions.

High battery's internal temperature is due to a high electric current including operations with quick charging and the discharging speed. Therefore, the appropriate management of the battery temperature is crucial to ensure efficient performance and safe functioning. In our approach, temperature monitoring is one of the fundamental management processes. However, monitoring the temperature distribution within the batteries is not straightforward. We adopted the thermal model proposed in [162] for monitoring the internal battery temperature. The internal battery temperature (T_p) at a period p is estimated according to the observed temperature at the surface of the battery T_{surf} and the ambient temperature T_{amb} as in equation 6.1:

$$T_p = T_{surf} \times \left(1 + \frac{R_{in}}{R_{out}}\right) - T_{amb} \times \frac{R_{in}}{R_{out}} \tag{6.1}$$

R_{in} and R_{out} represent the thermal resistance inside and outside the battery. The feasibility proof of this model is discussed in [162].

6.3/ Heating prediction model

The operating temperature of the battery is a determinant factor of its deterioration. To predict the impact of the node selection as a CH on its state of health, it is crucial to predict the effects of CH's activities on the battery temperature rise. In [163], Taheri and Bahrami studied the Lithium-Ion batteries temperature rise according to the activity degree. The results of this work show that, under a constant discharging intensity, LIB's temperature curves follow the same global scheme corresponding to three stages. When the battery is full charged (75%-100%) the temperature rises moderately. When the battery is averagely charged (10%-75%) the battery temperature rises slowly. Finally, when the battery is almost discharged, the temperature rises quickly. However the global rapidity of the temperature rise depends on the convective transfer coefficients of the battery surface h (the ability of the device to evacuate the heat) and the discharging intensity.

Based on the provided results in [163], we modeled the temperature rise when a node is a CH as in equation 6.2 :

$$T_{p+1} = T_p + \Delta T \times |p| \tag{6.2}$$

The values of ΔT are given in Kelvin per hour (K/h) and depends on the electrical intensity and the discharging stage as shown in Table 6.1. The presented values are those related to a Lithium Ion battery of $18mm \times 16mm \times 0.2mm$ size, a capacity of $0.27Wh$ (1000 joules) and a voltage of $3.7V$, and convective transfer coefficient h of $5w/(m^2K)$. The temperature T_{p+1} corresponds to the next period predicted temperature based of the measured temperature T_p computed using equation 6.1. $|P|$ corresponds to the time (in seconds) during which the node is CH.

Table 6.1: Temperature rising rate according to the discharging intensity and the charge level of the battery

Intensity	$SoC \in [75, 100]$	$SoC \in [10, 75]$	$SoC \in [0, 10]$
1.5A	2.6	1.07	8
3A	6.4	8.30	43
6A	40	36.9	140

6.4/ Contribution

In this section, we present a new distributed clustering technique based on the state of health prediction of rechargeable batteries called $LECA\_SOH$. CHs nodes perform more functionalities than CMs and accordingly deteriorate more quickly. The novelty of the proposed solution consists of predicting the degradation effect of a potential selection of the node as a CH. $LECA\_SOH$ selects as CHs the nodes that undergo less estimated battery deterioration during the forthcoming round, which in turn extends the network long-term lifetime. $LECA\_SOH$ computes a weight for each node, and the node holding the highest weight among its surrounding area is selected as a CH. The weight of a node i is computed using two metrics. First, the average neighborhood cardinality that depicts the number of neighbors within the communication range, $N_1(i)$. Indeed, a high node degree strengthens the cluster connectivity and reduces the intra-cluster communication. Secondly, the predicted battery state of health degradation (Δ_{SOH}) metric that estimates the potential degradation of the node battery if it is selected as CH during the subsequent round. The use of Δ_{SOH} allows preventing nodes that risk considerable damage in the following rounds from being chosen as CH. The computation of SOH_p is detailed in section 5.1.2. The future SOH state is then expressed by equation 6.3, where p refers to the current round.

$$SOH_{p+1}(i) = SOH_p(i) - \Delta SOH_p(i) \tag{6.3}$$

The weight W_i of a node i is computed based on the two previous metrics using equation 6.4, Where α and β depict the decision maker defined coefficients of the corresponding metrics.

$$W_i = \alpha \times |N_1(i)| + \beta \times \Delta_{SOH}(i) \wedge \alpha + \beta = 1 \tag{6.4}$$

6.4.1/ Weights computation

To compute the weights W_i at the beginning of a new round p, $LECA\_SOH$ assumes that each node i broadcasts a $Neighbor\_Discovery(ID_i)$ message to all surrounding nodes. A Neighbor j receiving this message, updates its local state by inserting node i information into $NeighborsList_j$. Next, node j replies by sending its local state $Local\_State_j(ID_j, W_j)$ to i. Progressively, node i knows the local state of its surrounding neighbors using the state list ($NeighborsList_i$). Every time, the node i updates its weight due to a received $Local\_State$ message from a neighbor, i broadcasts its new computed weight.

6.4.2/ CHs election phase

CHs election algorithm is inspired from the self-stabilization works [144], which allows $LECA\_SOH$ to adapt to transient events and network dynamic topology. We note CH_i the cluster head of the node i. $LECA\_SOH$ organizes each cluster into k-hop routing tree where the root is the cluster head node. The rooting tree defines how the data sensed by the devices are aggregated and forwarded to the CH. The CHs rely then the data to the BS using long range communication mode. The proposed approach is fully distributed where each node has only a partial vision of the network.

The clustering process of $LECA\_SOH$ is illustrated in figure 6.2. Whence the weights of the nodes are computed, each node checks if the conditions to be the cluster head during the next round are satisfied. The node c is the cluster head of i (c may equal to i) if and only if c is within a range of k hops from i and its weight is higher than the weights of all $N_{<k}(i)$. More formally:

$$CH_i = c \Leftrightarrow \begin{cases} c \in N_{<k}(i) \cup i \\ W_c \geq W_i \\ \forall j \in N(i) : W_c \geq W_j \\ \forall j \in N(i), CH_j \in N_{<k}(i) : \\ W_c \geq W_{CH_j} \end{cases} \tag{6.5}$$

If two or more nodes in $N_{<k}$ possess the same highest weight, the distance is used as

a tie-breaker (by choosing the nearest CH). Furthermore, if node i is isolated or no CH satisfies the multi-hop constraint, then node i elects itself as CH. After CHs election, each node exploits its direct neighbors knowledge to shape the multi-hop cluster topology. The nodes update their distance toward the CH and select their parents, P_i, in the routing tree by following the shortest path to the selected CH as expressed in following equation 6.6.

$$P_i = m \Leftrightarrow \begin{array}{l} \forall j \in N(i), Ch_i = CH_j : \\ HopsToCH_j \geq HopsToCH_m \end{array} \tag{6.6}$$

$HopsToCH_i$ refers to the minimum number of hops to reach the CH_i from i and it is measured using the recursive formula:

$$HopsToCH_i = \left(\min_{\substack{j \in N(i) \\ CH_i = CH_j}} HopsToCH_j \right) + 1 \tag{6.7}$$

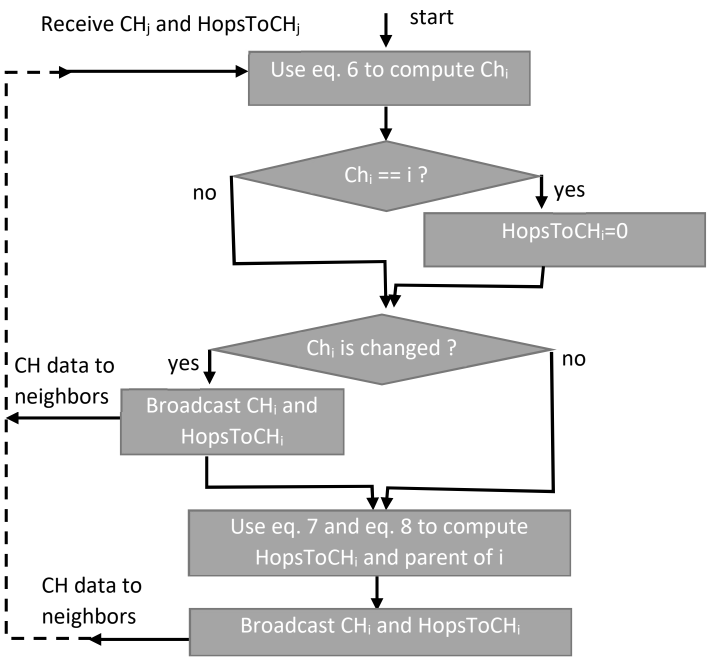


Figure 6.2: Flowchart of $LECA\_SOH$ leader election phase

$LECA\_SOH$ functioning is mainly depending on the ability to estimate the battery degradation if a node is CH. In this sections, we present the different models used for the implementation of our solution.

6.5/ Simulation

6.5.1/ Simulation parameters

In this experiment, we adopted the network parameters described in section 4.7.1. The experiment parameters are specified in Table 6.2. We assume that when a node is no longer a CH, its temperature gradually cools until reaching the ambient temperature (linear cooling). The performance of our clustering approach is evaluated by considering the single-hop and multi-hop (with K=2 hops) intra-clustering scenarios. The proposed scheme is tested with different network density to analyze the performance under different scenarios.

$LECA\_SOH$ is, to the best of our knowledge, the first clustering approach that considers the prediction of the battery SoH in the clustering process. To show the relevance of this new paradigm, the performances of the proposed scheme are compared with IEECP [19] and O-LEACH [90]. These protocols belong to the same clustering class and their primary goal is the energy efficiency and extending the network lifetime. Five parameters are considered for the analysis of the protocols performance, namely, the CHs cardinality, the energy consumption, the network lifetime (in short and long-term), the number of recharging cycles and the average number of dead nodes.

Table 6.2: Experiment setting parameters

Parameter	**Value**
Network size	$1000 \times 1000\ m^2$
Node density (δ)	$\delta \in [200, 1000]$
Transmitting range (Tr)	90 m
α, β	0.5, 0.5
E_{elec}	$50 mJ/bit$
ε_{FS}	$10\ nJ/bit/M^2$
ε_{Mfs}	$0.0013\ nJ/bit/M^4$
R_{in},R_{out}	$3.2KW^{-1}$,$8.44KW^{-1}$
Data packet size	$100\ bytes$
Battery's full charge	1 KJ

6.5.2/ Experimental Results

6.5.2.1/ CH's cardinality

Typically, a reduced CH's cardinality reveals efficient clustering performance as it depicts the number of wireless communication channels formed with the BS and the number of clusters generated. Figure 6.3 illustrates the average number of CHs of the proposed approach versus O-LEACH and IEECP protocols under different density values $\delta \in [200, 1000]$. For the case of a single hop transmission (Figure 6.3), the number of clusters generated by the proposed scheme is stabilized between 70 and 95 clusters. Whereas, this value increases up to 105 and 115 with IEECP and O-LEACH respectively. Figure 6.3 also illustrates the case when the multi-hop scheme is considered. In general, the clusters cardinality decreases. The average CHs generated is reduced by an average of 31.4% as compared to the single-hop scheme, the CHs cardinality in this case varies between [40-55]. This can be explained by the fact that clusters coverage area increases, in multi-hop scheme, which allows the clusters to manage more nodes. $LECA\_SOH$ maintains the lowest CH cardinality, it shows an enhancement of 10.3% and 6.1% compared to O-LEACH and IEECP respectively. This improvement is due to the consideration of the nodes connectivity metric in the clustering process of $LECA\_SOH$, which leads to the election of CHs with high connectivity. Therefore, more nodes are covered by a single cluster.

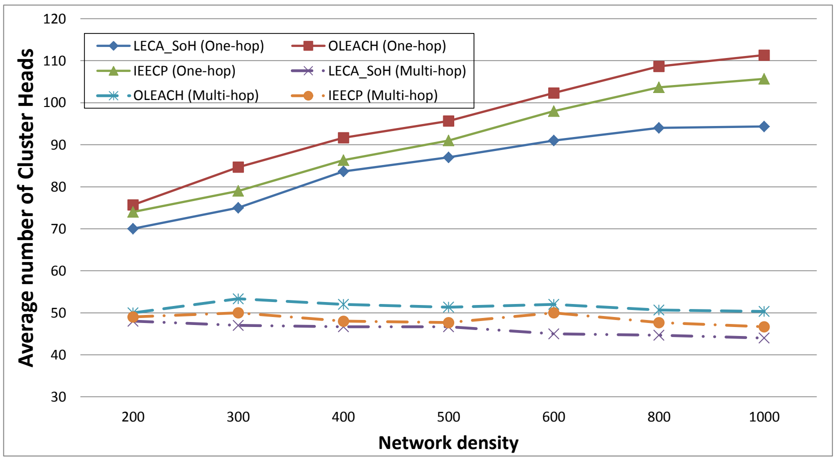


Figure 6.3: Average Cluster Head cardinality considering the Single-hop clustering and the Multi-hop clustering.

6.5.2.2/ Average energy consumed

Figure 6.4 exhibits the average energy consumed by network nodes by considering a variable network density $\delta \in [200 - 1000]$ and by assuming the single-hop and multi-hop scenario. From these figures, we observe that the energy required for the clustering is affected by the network density and the clustering scheme used. When the number of nodes increases, the energy required to build the cluster structure increases as well. This value is bounded by 0.95×10^{-3} KJoule when the single hop scenario is considered, while it reaches up to 1.4×10^{-3} KJoule when the multi-hop clustering is employed. This is due to the amount of messages exchanged to construct the structure. According to figure 6.4, IEECP consumes the highest amount of energy compared to the other approaches owing to the modified fuzzy C-means algorithm used during the clustering which requires the exchange of additional setup packets to form the clusters. Accordingly, O-LEACH consumes 39% less energy compared to IEECP.

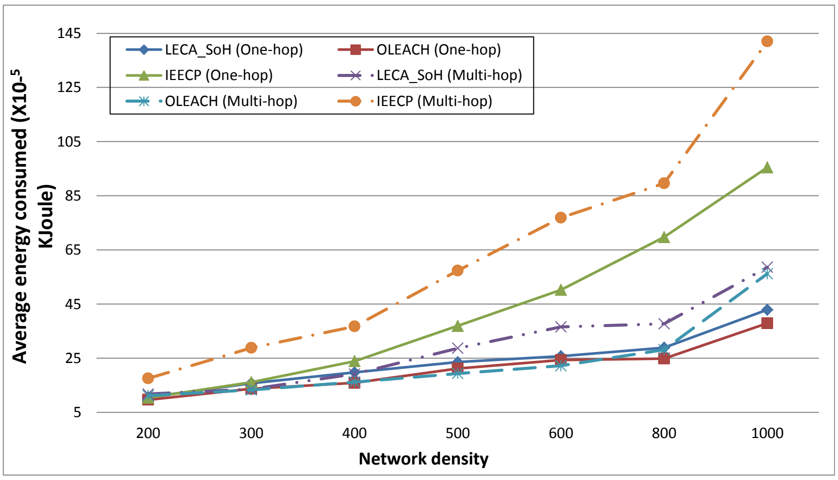


Figure 6.4: Average energy consumed with the single-hop and the Multi-hop clustering.

$LECA\_SOH$ shows an average improvement of 27.3% compared to IEECP. O-LEACH consumes the lowest energy. It consumes slightly less energy than $LECA\_SOH$ (11% less) because this later requires more messages to compute the weight metric (computed using Equation 6.4), which consumes an extra more energy. However, the result is acceptable because the topology generated by the proposed approach carries less clusters (as illustrated in previous section 6.5.2.1) and exhibits a better lifetime performance in the long-term (which is discussed in section 6.5.2.4 and 6.5.2.5).

6.5.2.3/ Network lifetime (in the short-term)

This measure stands for the average number of rounds elapsed before the first network device spend all of it energy and its battery needs recharging. This measure is typically used in the literature to evaluate the energy consumption of devices. It also depicts the lifespan of one battery cycle by considering only the residual energy (short-term vision).

Figure 6.5 illustrates the average battery cycle lifetime of the studied approaches by considering the single hop and multi hop clustering modes. From this figure, we observe that the network lifetime increases when the multi-hop is considered, as compared to the single-hop clustering. The battery cycle lifetime varies between [900 – 2600] rounds in the single-hop scenario and it increases to [1200 – 2800] rounds in the multi-hop scenario. Indeed, as the network scales, ensuring the connectivity tends to be more difficult.

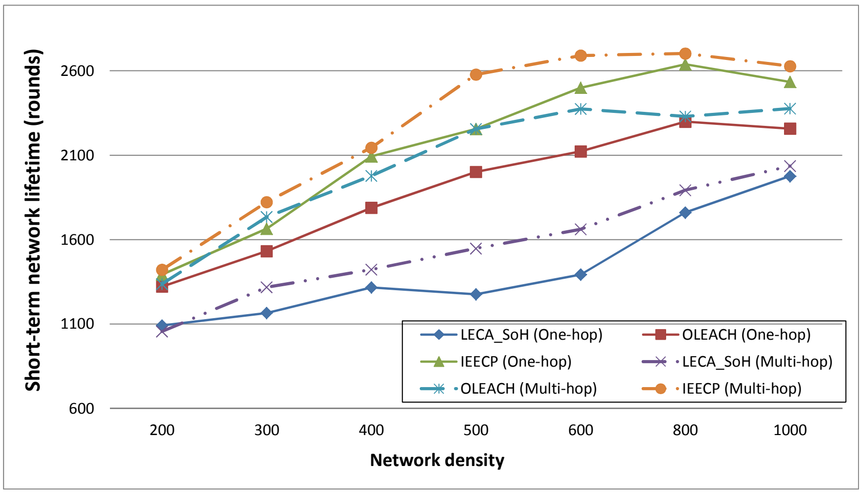


Figure 6.5: Average network lifetime (in the short-term) in single-hop and multi-hop scenario.

Therefore, multi-hop communications are advantageous to reduce the energy cost devoted to wireless transmission. Moreover, we observe that the network lifespan, for the three schemes, increases when the network is more connected. Indeed, when the density increases, the clusters connectivity increases and the CHs are able to cover more nodes (Figure 6.3), which in turn reduces the long range transmission and boosts the energy preservation.

Based on the shape of figure 6.5, we observe that IEECP maintains the highest short-term durability (without battery recharging). This figure shows an average improvement of 9% compared to O-LEACH. This improvement is due to the consideration of communication reliability in IEECP in addition to the residual energy of nodes. Both O-LEACH and IEECP show an average improvement of 19.7% compared to our approach. This is because O-LEACH and IEECP consider the residual energy of devices, which results in a longer short-term lifetime (over one battery cycle only). In the next paragraph we study the performances of the three approaches according to the long-term durability (the batteries life over several recharging cycles).

6.5.2.4/ Network lifetime (in the long-term)

This metric depicts the average number of rounds elapsed before the appearance of the first completely dead node in the network (i.e. the SoH of its rechargeable battery is totally exhausted and needs to be replaced). This metric enables the evaluation of the network durability and estimates the devices battery degradation aspect in the long-term. Usually, the network lifetime is described in different manners, e.g. the interval from network initialization to the last node death round (LND) [89] or the interval from the initial deployment until a predefined percentage of nodes death. In this experiment, we consider the network lifetime as the period until the first node death (FND).

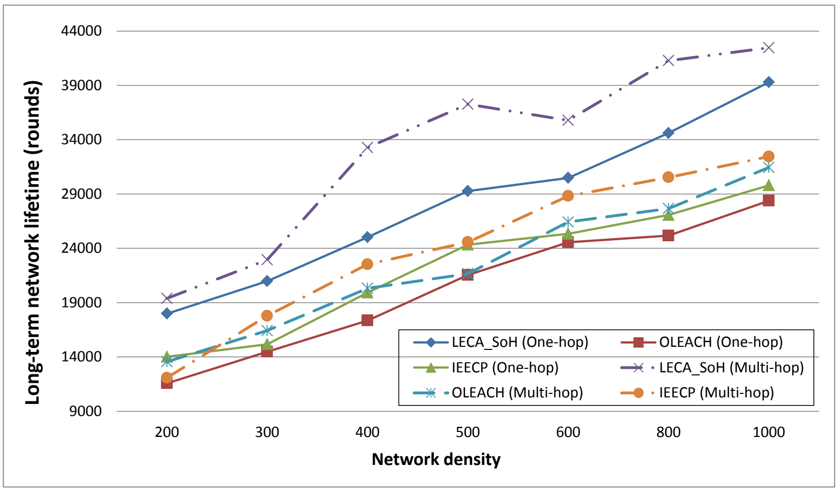


Figure 6.6: Average network lifetime (long-term) in single-hop and multi-hop clustering.

Figure 6.6 shows the average network lifetime in long-term of our approach versus IEECP and O-LEACH. Initially, nodes tend to be isolated due to the low connectivity, they are self elected as CH to send their data toward the BS. Hence, at this point, the lifetime of the three protocols is relatively short. When the density increases, nodes connectivity increases, which in turn lowers the CHs cardinality and balances the consumption of energy within CMs. From figure 6.6, we observe that nodes using O-LEACH tend to die earlier and the FND only reached 28400 rounds (with $\delta = 1000$) compared to IEECP where the nodes survive up to 29800 rounds. This figure shows that our proposed approach exhibits better performances, the FND node in $LECA\_SOH$ reached 32500 rounds. This improvement is due to the use of the prediction metric Δ_{SOH}. Indeed, instead of performing the clustering based on the actual state of charge, $LECA\_SOH$ predicts the amount of battery degradation for the following round before the election of the potential CH's. Accordingly, inadequate nodes are avoided and the battery aging process is delayed, which strengthens the network durability. Moreover, we observe that considering the SoH prediction in the clustering process balances the temperature fairly among devices, which further reduces the battery degradation. Figure 6.6 illustrates a comparison of the three approaches by considering the multi-hop clustering. $LECA\_SOH$ shows better lifetime performances compared to the classic approaches in both single-hop and multi-hop scenarios (improved by 23% and up to 44% respectively).

6.5.2.5/ Average number of recharging cycles and dead nodes (long-term)

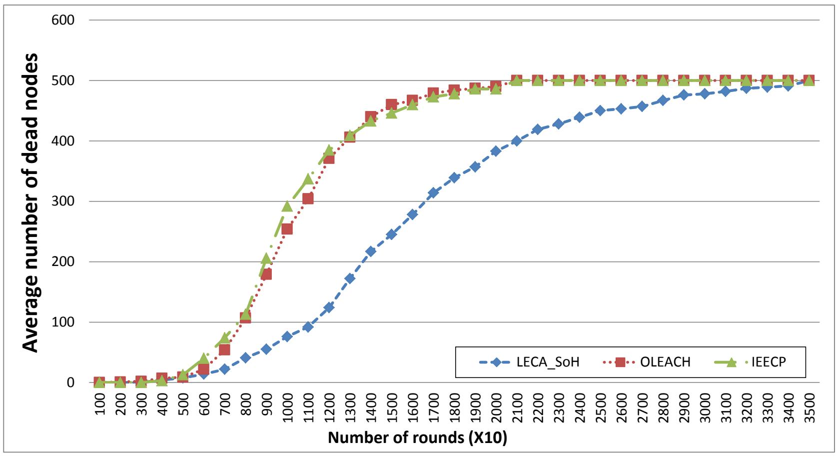


Figure 6.7: Average number of dead nodes (δ = 500 nodes)

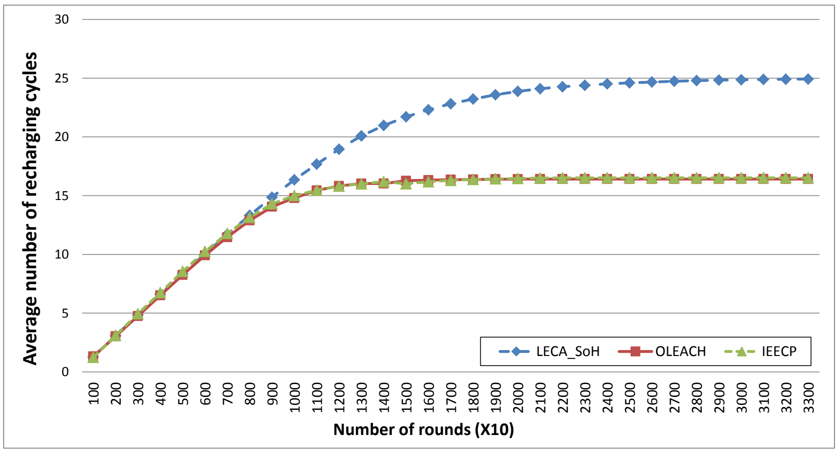


Figure 6.8: Average number of battery recharging cycles

The average number of recharging cycles and the average number of completely dead nodes (i.e. nodes with a completely damaged battery, $SOH = 0$) represent two significant parameters to demonstrate the lifetime efficiency of the proposed approach. These parameters depict the elapsed period before a rechargeable battery is entirely dead and needs to be changed, i.e. it provides an estimation of the material substitution frequency.

Figures 6.7 and 6.8 illustrate the average number of completely dead nodes and the number of battery recharging cycles (measured in rounds) of $LECA\_SOH$ vs O-LEACH and IEECP with $\delta = 500$ nodes. It can be observed from the shape of figure 6.8 that O-LEACH and IEECP withstand approximately the same number of cycle (up to 16 cycles). Whereas, in $LECA\_SOH$ the average number of recharging cycle reached 25 cycles. This lifespan improvement is mainly due to the use of the SoH prediction metric in the clustering process. Therefore, the number of completely dead nodes is considerably reduced as shown in figure 6.7. The proposed approach improves the number of recharging cycles by an average of 31.4% and reduces the average number of nodes death by 34.6% when compared to the conventional energy efficient approaches [90; 19].

6.6/ Conclusion

In this chapter, we proposed an energy efficient prediction based clustering approach called $LECA\_SOH$. The approach is distributed and aims to optimize the network lifespan in long-term by considering the battery degradation aspect in the decision making process of the wireless networking protocols. More precisely, $LECA\_SOH$ uses an SoH prediction based mechanism during the CHs election in order to elect nodes that will endure less potential degradation during the forthcoming rounds, which in turn extends the system lifetime. The objective of the contribution is the integration of the battery degradation prediction in the clustering process to extend the network lifespan in the long-term. Moreover, the approach contemplates the behavior of the rechargeable batteries and considers the thermal effect on the battery degradation. The obtained results are encouraging as the lifespan of the network is extended with the use of the prediction criterion. The aging based approach outperforms conventional methods in terms of clusters cardinality and network lifetime in the long-term (exhibits an average improvement of 44%).

General conclusion

The emerging paradigm of the Internet of Things (IoT) is rapidly making its path across our modern life. IoT networks aims to improve the quality of our life's by enabling the connection of many smart devices, technologies, and applications. This paradigm becomes one of the most active research topics and attracted the attention of researchers from academic and industrial sectors due to its potential benefits. Despite the valuable services that these paradigm cover, IoT networks have many design challenges and limitations (e.g. devices restrained resource) that should be treated by researchers, otherwise, the quality of services in IoT applications will negatively deteriorate (scalability, energy efficiency and fault tolerance issues, etc.).

Preserving the devices' energy, increasing the network lifetime, decreasing communication latency and improving the fault tolerance are the primary concerns that should be considered when building IoT applications. Currently, IoT landscape is characterized by the deployment of a vast number of heterogeneous smart devices with a restrained storage capacity, processing and energy, communicating through an error prone wireless links. Due to that limitation, an efficient management of communication channels becomes important in order to improve the network performance. In addition, when the vision of the upcoming generation of IoT networks will be fully accomplished, the scalability and maintenance cost will be enormous. Disregarding these challenges will restrain the performance of the IoT solutions. The majority of IoT elements are rechargeable battery-powered devices which deteriorate over time and require maintainability. Although the battery itself might be inexpensive, its substitution implies a costly labor. Accordingly, the topology management and the choice of an efficient network structuring technique is a pertinent topic, especially with the expanding behavior and the dynamicity of these networks. Moreover, the effective management of devices batteries needs to be applied to maintain network longevity.

In this thesis we addressed the management problem of the shared communication channel, the management of the energy resource in long term and the management of the IoT networks topology in general. In this context, Clustering is one of the relevant techniques for structuring the network, which aims to optimize the network performance and scalability. In this thesis, we explored this issue to propose new clustering algorithms suitable for the IoT structure. We firstly presented general information on IoT networks and the layered model providing communications in such networks. We presented background knowledge on cross-layer approaches between the routing and the MAC layer and we presented some existing work in this research axis. Then, we presented taxonomy on hierarchical clustering algorithm considering relevant key parameters. Based on the limitation of existing management solutions for IoT networks we proposed in the third chapter different contributions suitable for IoT network. The common objective of these approaches is to optimize the network performance and improve the network durability.

In the first contribution, we provided a distributed cross layer TDMA scheduling algorithm aware routing information (WB-DTSA). The approach aims to manage the access to the communication channel by using the surrounding topology knowledge to perform the transmitting slot allocation. Time slot are allocated to nodes depending on their position in the routing tree while privileging nodes with long routes to the base station by allocating them earlier transmitting slots to reduce the latency. The approach considers the average distance and the interference degree among each node and its neighborhood to further improve the slot allocation efficiency and reduce the energy wastage.

In the second contribution, we suggested DC2HC, a distributed approach for k-hop intra-clustering for large networks. The approach optimizes the set of representative cluster heads and extends the network lifetime. The cluster heads election is performed based on the two-hop topology knowledge of nodes to strength the cluster connectivity. It also considers the residual energy of nodes to balance the energy consumption among larger clusters. The contribution has a linear time and space complexity and provides the convergence proof of the proposed solution.

Afterward, we presented some initial works on how the battery degradation level could be considered in the decision making process of the wireless networking protocols. The purpose is to optimize the devices lifetime in order to diminish the material alternation frequency in order to impact the production and the distribution load. Indeed, the scale down of production and distribution load is the best way to reduce the carbon footprint of the ITC, (i.e long-term vision of energy optimization). We considered the rechargeable aspect of devices batteries and we discuss how a long-term vision of energy consumption can be conducted by taking into account the devices state of health (SOH).

The third contribution, we presented a multi-hop clustering technique based on a long-term vision of energy optimization (LTEOC) by considering the battery aging aspect of rechargeable batteries in a multi-hop clustering aspect. The proposed scheme enhances the battery behavior and extends the lifespan of network devices in the long-term, which has a positive impact on green computing environment.

In the last proposal, we conceived a distributed SOH prediction based Clustering Approach for Long-term Energy optimization ($LECA\_SOH$). This latter uses an original metric which predict the forthcoming SOH of devices battery before the clusters formation. As CHs nodes perform more functionality than CMs and accordingly deteriorate more quickly. The novelty of the proposed solution consists of predicting the degradation effect of a potential selection of the node as a CH.

The experimental results demonstrate that the proposed approaches performs better compared to other referenced approaches in terms of different performance parameters, such as Communication latency, schedule length, cluster heads cardinality, messages overhead, network lifetime in short and long-term.

6.7/ FUTURE WORK AND PERSPECTIVES

Although the proposed solutions for the management of IoT networks presented in this thesis have led to encouraging results and acceptable performance. There are other significant topics that we plan to address and study further in the research for the next steps. We summarize some interesting future perspectives as follow:

The consideration of mobile multi-sinks is appealing to enable the tolerance of more dynamic topologies since mobile sinks move continuously over the network in a more or less random fashion. Moreover, mobile devices make it possible to expand the network coverage and improve the monitoring of the physical phenomenon. Indeed, the mobility can significantly improve the overall network lifetime, reliable data gathering, delay and latency problems [164]. However, with a mobile multi-sinks approach, several aspects need to be exploited with the addition of mobility, e.g. network self-organization techniques used to enable the organization of the network object.

To highlight the strengths and weaknesses of our protocols, we plan to conduct some experiments by using a real physical wireless network under various deployments and sensing environments. We are convinced of the need to extend our simulations to other communication models. The UDG model is known to be idealistic, however, does not reflect the actual communication model of modern devices. The usage of a realistic communication model would allow performance analysis of our different approaches in a less idealistic context. Moreover, experiment using other network topologies, e.g. urban areas guided by different placement for the based station and the presence of different obstacles seem to be interesting in order to observe the performance of our approaches under different scenarios.

The machine learning is an interesting topic to focus on, we aim to integrate this technique to predict the energy consumption of nodes in long term and use this information to perform an optimal energy routing process to further preserve the energy consumption and increase the devices lifespan.

We also aim to use the machine learning to perform an optimal transition scheduling. Indeed, with machine learning it is possible to extract knowledge from the network and progressive learning in the presence of inherent uncertainties and the lack of network state information. The routing mechanism can use machine learning to learn and to predict the devices behavior and the availability of the channels. In this scenario, it is possible to dynamically choose the interval of exchanging messages and the optimal routing paths after the learning period. This consideration aims to considerably reduce the network overhead and preserve more energy, which in turn further extend the network lifetime in long-term. Machine learning can also be used to optimize other factors like resource usage, estimation of response times and data traffic monitoring.

Bibliography

[1] M. S. Batta, H. Mabed, Z. Aliouat, S. Harous, Battery state-of-health prediction based clustering for lifetime optimization in iot networks, IEEE Internet of Things Journal.

[2] M. S. Batta, H. Mabed, Z. Aliouat, S. Harous, A distributed multi-hop intra-clustering approach based on neighbors two-hop connectivity for iot networks, Sensors 21 (3) (2021) 873.

[3] M. S. Batta, Z. Aliouat, H. Mabed, M. Merah, An improved lifetime optimization clustering using kruskal's mst and batteries aging for iot networks, in: 2022 International Symposium on Networks, Computers and Communications (ISNCC), IEEE, 2022, pp. 1–6.

[4] M. S. BATTA, H. MABED, Z. ALIOUAT, State of health optimization based unequal clustering in iot networks, in: 2022 International Wireless Communications and Mobile Computing (IWCMC), IEEE, 2022, pp. 125–130.

[5] M. S. BATTA, Z. ALIOUAT, H. MABED, M. MERAH, A distributed energy-efficient unequal clustering based kruskal heuristic for iot networks, in: 14th International Conference on Applications of Graph Theory in Wireless Ad hoc Networks and Sensor Networks (GRAPH-HOC), Vol. 12, 2022, pp. 117–129.

[6] M. Merah, Z. Aliouat, M. S. Batta, A hybrid neural network and graph theory based clustering protocol for dynamic iot networks, in: 2022 International Conference on Advanced Aspects of Software Engineering (ICAASE), IEEE, 2022, pp. 1–7.

[7] H. Mabed, M. Batta, Z. Aliouat, Optimization of rechargeable battery lifespan in wireless networking protocols, in: MobiQuitous 2020-17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, 2020, pp. 118–124.

[8] M. S. Batta, Z. Aliouat, H. Mabed, S. Harous, Lteoc: Long term energy optimization clustering for dynamic iot networks, in: IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 2020.

[9] M. S. Batta, H. Mabed, Z. Aliouat, Dynamic clustering based energy optimization for iot network, in: International Symposium on Modelling and Implementation of Complex Systems, Springer, 2020, pp. 77–91.

[10] M. S. Batta, Z. Aliouat, S. Harous, A distributed weight-based tdma scheduling algorithm for latency improvement in iot, in: IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 2019, pp. 0768–0774.

[11] M. S. Batta, S. Harous, L. Louail, Z. Aliouat, A distributed tdma scheduling algorithm for latency minimization in internet of things, in: 2019 IEEE International Conference on Electro Information Technology (EIT), 2019, pp. 108–113.

[12] L. Vailshery, Iot and non-iot connections worldwide 2010-2025, https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/ (Statista, 2018).

[13] A. Alkhatib, G. Baicher, Wireless sensor network architecture, in: International conference on computer networks and communication systems (CNCS 2012), IACSIT Press Singapore, 2012.

[14] N. Baccour, A. Kouba, L. Mottola, M. A. Zuniga, H. Youssef, C. A. Boano, M. Alves, Radio link quality estimation in wireless sensor networks: A survey, ACM Transactions on Sensor Networks (TOSN) 8 (4) (2012) 1–33.

[15] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, Internet of things: A survey on enabling technologies, protocols, and applications, IEEE communications surveys & tutorials 17 (4) (2015) 2347–2376.

[16] K. Aldahdouh, K. Darabkh, W. Al-Sit, et al., A survey of 5g emerging wireless technologies featuring lorawan, sigfox, nb-iot and lte-m, in: International conference on wireless communications signal processing and networking (WiSPNET), IEEE, 2019, pp. 561–566.

[17] L. Atzori, A. Iera, G. Morabito, The internet of things: A survey, Computer networks 54 (15) (2010) 2787–2805.

[18] P. Ray, A survey on internet of things architectures, Journal of King Saud University Computer and Information Sciences 30 (3) (2018) 291–319.

[19] A. Hassan, W. Shah, A. Habeb, M. Othman, M. Al-Mhiqani, An improved energy-efficient clustering protocol to prolong the lifetime of the wsn-based iot, IEEE Access 8 (2020) 200500–200517.

[20] P. Sethi, S. Sarangi, Internet of things: Architectures, protocols, and applications, Journal of Electrical and Computer Engineering 2017 (2017) 1–25.

[21] R. A. Sadek, Hybrid energy aware clustered protocol for iot heterogeneous network, Future Computing and Informatics Journal 3 (2) (2018) 166–177.

[22] S. Hammoudi, Z. Aliouat, S. Harous, Challenges and research directions for internet of things, Telecommunication Systems 67 (2) (2018) 367–385.

[23] A. Shahraki, A. Taherkordi, o. Haugen, F. Eliassen, Clustering objectives in wireless sensor networks: A survey and research direction analysis, Computer Networks 180 (2020) 107376.

[24] E. Hossain, M. Rasti, L. Le, Radio resource management in wireless networks: an engineering approach, Cambridge University Press, 2017.

[25] S. Sinche, D. Raposo, N. Armando, A. Rodrigues, F. Boavida, V. Pereira, J. Silva, A survey of iot management protocols and frameworks, IEEE Communications Surveys & Tutorials 22 (2) (2019) 1168–1190.

[26] S. Dehkordi, K. Farajzadeh, J. Rezazadeh, R. Farahbakhsh, K. Sandrasegaran, M. Dehkordi, A survey on data aggregation techniques in iot sensor networks, Wireless Networks 26 (2) (2020) 1243–1263.

[27] P. Mishra, S. Verma, A survey on clustering in wireless sensor network, in: International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, 2020, pp. 1–5.

[28] S. Arjunan, S. Pothula, A survey on unequal clustering protocols in wireless sensor networks, Journal of King Saud University Computer and Information Sciences 31 (3) (2019) 304–317.

[29] V. Sucasas, A. Radwan, H. Marques, J. Rodriguez, S. Vahid, R. Tafazolli, A survey on clustering techniques for cooperative wireless networks, Ad Hoc Networks 47 (2016) 53–81.

[30] R. Mohamed, A. Saleh, M. Abdelrazzak, A. Samra, Survey on wireless sensor network applications and energy efficient routing protocols, Wireless Personal Communications 101 (2) (2018) 1019–1055.

[31] R. E. Mohamed, A. I. Saleh, M. Abdelrazzak, A. Samra, Survey on wireless sensor network applications and energy efficient routing protocols, Wireless Personal Communications 101 (2) (2018) 1019–1055.

[32] L. Xu, R. Collier, G. M. O'Hare, A survey of clustering techniques in wsns and consideration of the challenges of applying such to 5g iot scenarios, IEEE Internet of Things Journal 4 (5) (2017) 1229–1249.

[33] R. M. Curry, J. Smith, A survey of optimization algorithms for wireless sensor network lifetime maximization, Computers & Industrial Engineering 101 (2016) 145–166.

[34] H. Zimmermann, Osi reference model-the iso model of architecture for open systems interconnection, IEEE Transactions on communications 28 (4) (1980) 425–432.

[35] A. Sgora, D. Vergados, D. Vergados, A survey of tdma scheduling schemes in wireless multihop networks, ACM Computing Surveys (CSUR) 47 (3) (2015) 1–39.

[36] S. Li, L. Da Xu, S. Zhao, The internet of things: a survey, Information Systems Frontiers 17 (2) (2015) 243–259.

[37] J. Gubbi, R. Buyya, S. Marusic, M. Palaniswami, Internet of things (iot): A vision, architectural elements, and future directions, Future generation computer systems 29 (7) (2013) 1645–1660.

[38] F. Karray, M. W. Jmal, A. Garcia-Ortiz, A. Abid, M.and Obeid, A comprehensive survey on wireless sensor node hardware platforms, Computer Networks 144 (2018) 89–110.

[39] S. Kiran, G. Gupta, A pragmatic analysis on wireless sensors and actuator technologies for security in internet of things, International Journal of Future Generation Communication and Networking 13 (4) (2020) 5109–5116.

[40] L. Xu, G. M. O'Hare, R. Collier, A smart and balanced energy-efficient multihop clustering algorithm (smart-beem) for mimo iot systems in future networks, Sensors 17 (7) (2017) 1574.

[41] S. M. Darroudi, C. G., Bluetooth low energy mesh networks: A survey, Sensors 17 (7) (2017) 1467.

[42] S. Meka, B. Fonseca, Improving route selections in zigbee wireless sensor networks, Sensors 20 (1) (2020) 164.

[43] Y. Sung, S. Lee, M. Lee, A multi-hop clustering mechanism for scalable iot networks, Sensors 18 (4) (2018) 961.

[44] M. Chiani, A. Elzanaty, On the lora modulation for iot: Waveform properties and spectral analysis, IEEE Internet of Things Journal 6 (5) (2019) 8463–8470.

[45] R. Molina-Masegosa, J. Gozalvez, M. Sepulcre, Comparison of ieee 802.11 p and lte-v2x: An evaluation with periodic and aperiodic messages of constant and variable size, IEEE Access 8 (2020) 121526–121548.

[46] M. Bouaziz, A. Rachedi, A survey on mobility management protocols in wireless sensor networks based on 6lowpan technology, Computer Communications 74 (2016) 3–15.

[47] A. Maass, D. Nevsic, R. Postoyan, P. Dower, Lp stability of networked control systems implemented on wirelesshart, Automatica 109 (2019) 108514.

[48] M. B. Yassein, W. Mardini, A. Khalil, Smart homes automation using z-wave protocol, in: 2016 International Conference on Engineering & MIS (ICEMIS), 2016, pp. 1–6.

[49] S. Kaur, R. N. Mir, Clustering in wireless sensor networks-a survey, International Journal of Computer Network and Information Security 8 (6) (2016) 38.

[50] M. Lauridsen, I. Kovacs, P. Mogensen, M. Sorensen, S. Holst, Coverage and capacity analysis of lte-m and nb-iot in a rural area, in: 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall), 2016, pp. 1–5.

[51] R. Sinha, Y. Wei, S. Hwang, A survey on lpwa technology: Lora and nb-iot, Ict Express 3 (1) (2017) 14–21.

[52] R. Dangi, P. Lalwani, G. Choudhary, I. You, G. Pau, Study and investigation on 5g technology: A systematic review, Sensors 22 (1) (2022) 26.

[53] M. Mohammed Nasr, A. Abdelgader, Z. Wang, L. Shen, Vanet clustering based routing protocol suitable for deserts, Sensors 16 (4) (2016) 478.

[54] C. Nay, Sensors remind doctors to wash up, in: IBM Research Blog, 2013.

[55] J. Stankovic, Research directions for the internet of things, IEEE internet of things journal 1 (1) (2014) 3–9.

[56] S. Chen, H. Xu, D. Liu, B. Hu, H. Wang, A vision of iot: Applications, challenges, and opportunities with china perspective, IEEE Internet of Things journal 1 (4) (2014) 349–359.

[57] S. Chouikhi, I. El Korbi, Y. Ghamri-Doudane, L. Saidane, A survey on fault tolerance in small and large scale wireless sensor networks, Computer Communications 69 (2015) 22–37.

[58] J. Granjal, E. Monteiro, J. Silva, Security for the internet of things: a survey of existing protocols and open research issues, IEEE Communications Surveys & Tutorials 17 (3) (2015) 1294–1312.

[59] Y. Harbi, Z. Aliouat, S. Harous, A. Bentaleb, A. Refoufi, A review of security in internet of things, Wireless Personal Communications 108 (1) (2019) 325–344.

[60] D. Macedo, L. Guedes, I. Silva, A dependability evaluation for internet of things incorporating redundancy aspects, in: 11th IEEE international conference on networking, sensing and control, 2014, pp. 417–422.

[61] N. Maalel, E. Natalizio, A. Bouabdallah, P. Roux, M. Kellil, Reliability for emergency applications in internet of things, in: IEEE international conference on distributed computing in sensor systems, 2013, pp. 361–366.

[62] C. Sarkar, S. Nambi, R. Prasad, A. Rahim, A scalable distributed architecture towards unifying iot applications, in: World Forum on Internet of Things (WF-IoT), IEEE, 2014, pp. 508–513.

[63] F. Li, M. Vögler, M. Claessens, S. Dustdar, Efficient and scalable iot service delivery on cloud, in: Sixth international conference on cloud computing, IEEE, 2013, pp. 740–747.

[64] I. Silva, R. Leandro, D. Macedo, L. Guedes, A dependability evaluation tool for the internet of things, Computers & Electrical Engineering 39 (7) (2013) 2005–2018.

[65] A. Botta, W. De Donato, V. Persico, A. Pescape, Integration of cloud computing and internet of things: a survey, Future generation computer systems 56 (2016) 684–700.

[66] L. Song, K. K. Chai, Y. Chen, J. Schormans, J. Loo, A. Vinel, Qos-aware energy-efficient cooperative scheme for cluster-based iot systems, IEEE Systems Journal 11 (3) (2017) 1447–1455.

[67] C. Stefanovic, P. Popovski, D. Vukobratovic, Frameless aloha protocol for wireless networks, IEEE Communications Letters 16 (12) (2012) 2087–2090.

[68] E. Ziouva, T. Antonakopoulos, Csma/ca performance under high traffic conditions: throughput and delay analysis, Computer communications 25 (3) (2002) 313–321.

[69] A. Tsertou, D. Laurenson, Revisiting the hidden terminal problem in a csma/ca wireless network, IEEE Transactions on Mobile Computing 7 (7) (2008) 817–831.

[70] Y. Li, X. Zhang, J. Zeng, Y. Wan, F. Ma, A distributed tdma scheduling algorithm based on energy-topology factor in internet of things, IEEE Access 5 (2017) 10757–10768.

[71] M. Zanjireh, H. Larijani, A survey on centralised and distributed clustering routing algorithms for wsns, in: 2015 IEEE 81st Vehicular Technology Conference (VTC Spring), 2015, pp. 1–6.

[72] I. Chakeres, E. Belding-Royer, Aodv routing protocol implementation design, in: 24th International Conference on Distributed Computing Systems Workshops, 2004. Proceedings., IEEE, 2004, pp. 698–703.

[73] J. Kulik, W. Heinzelman, H. Balakrishnan, Negotiation-based protocols for disseminating information in wireless sensor networks, Wireless networks 8 (2) (2002) 169–185.

[74] K. Akkaya, M. Younis, A survey on routing protocols for wireless sensor networks, Ad hoc networks 3 (3) (2005) 325–349.

[75] G. Sara, D. Sridharan, Routing in mobile wireless sensor network: A survey, Telecommunication Systems 57 (1) (2014) 51–79.

[76] D. Baker, A. Ephremides, The architectural organization of a mobile radio network via a distributed algorithm, IEEE Transactions on communications 29 (11) (1981) 1694–1701.

[77] M. Gerla, J. Tzu-Chieh Tsai, Multicluster, mobile, multimedia radio network, Wireless networks 1 (3) (1995) 255–265.

[78] C. Chiang, M. Gerla, L. Zhang, Forwarding group multicast protocol (fgmp) for multihop, mobile wireless networks, Cluster Computing 1 (2) (1998) 187–196.

[79] P. Basu, N. Khan, T. Little, A mobility based metric for clustering in mobile ad hoc networks, in: Proceedings 21st international conference on distributed computing systems workshops, IEEE, 2001, pp. 413–418.

[80] A. Amis, R. Prakash, T. Vuong, D. Huynh, Max-min d-cluster formation in wireless ad hoc networks, in: Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), Vol. 1, 2000, pp. 32–41.

[81] M. Khan, F. Aadil, M. Maqsood, S. Bukhari, M. Hussain, Y. Nam, Moth flame clustering algorithm for internet of vehicle (mfca-iov), IEEE Access 7 (2018) 11613–11629.

[82] F. Aadil, W. Ahsan, Z. Rehman, P. Shah, S. Rho, I. Mehmood, Clustering algorithm for internet of vehicles (iov) based on dragonfly optimizer (cavdo), The Journal of Supercomputing 74 (9) (2018) 4542–4567.

[83] M. Ren, L. Khoukhi, H. Labiod, J. Zhang, V. Veque, A mobility-based scheme for dynamic clustering in vehicular ad-hoc networks (vanets), Vehicular Communications 9 (2017) 233–241.

[84] K. Liu, J. Su, J. Zhang, F. Liu, C. Gong, A novel stable cluster protocol for mobile ad hoc networks, in: 2005 IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, Vol. 2, IEEE, 2005, pp. 1328–1332.

[85] B. Hassanabadi, C. Shea, L. Zhang, S. Valaee, Clustering in vehicular ad hoc networks using affinity propagation, Ad Hoc Networks 13 (2014) 535–548.

[86] Z. Zhang, A. Boukerche, R. Pazzi, A novel multi-hop clustering scheme for vehicular ad-hoc networks, in: Proceedings of the 9th ACM international symposium on Mobility management and wireless access, 2011, pp. 19–26.

[87] L. Conceiccao, M. Curado, Onto scalable wireless ad hoc networks: Adaptive and location-aware clustering, Ad hoc networks 11 (8) (2013) 2484–2499.

[88] W. B. Heinzelman, A. Chandrakasan, H. Balakrishnan, An application-specific protocol architecture for wireless microsensor networks, IEEE Transactions on wireless communications 1 (4) (2002) 660–670.

[89] A. Ammar, A. Dziri, M. Terre, H. Youssef, Multi-hop leach based cross-layer design for large scale wireless sensor networks, in: International wireless communications and mobile computing conference (IWCMC), IEEE, 2016, pp. 763–768.

[90] J. Bhola, S. Soni, G. K. Cheema, Genetic algorithm based optimized leach protocol for energy efficient wireless sensor networks, Journal of Ambient Intelligence and Humanized Computing 11 (3) (2020) 1281–1288.

[91] T. Behera, S. Mohapatra, U. Samal, M. Khan, M. Daneshmand, A. Gandomi, Residual energy-based cluster-head selection in wsns for iot application, IEEE Internet of Things Journal 6 (3) (2019) 5132–5139.

[92] S. Singh, P. Kumar, J. Singh, A survey on successors of leach protocol, Ieee Access 5 (2017) 4298–4328.

[93] A. Hamzah, M. Shurman, O. Al-Jarrah, E. Taqieddin, Energy-efficient fuzzy-logic-based clustering technique for hierarchical routing protocols in wireless sensor networks, Sensors 19 (3) (2019) 561.

[94] J. Zhang, X. Feng, Z. Liu, A grid-based clustering algorithm via load analysis for industrial internet of things, IEEE Access 6 (2018) 13117–13128.

[95] H. Lin, D. Bai, D. Gao, Y. Liu, Maximum data collection rate routing protocol based on topology control for rechargeable wireless sensor networks, Sensors 16 (8) (2016) 1201.

[96] Y. Yang, C. Wang, Wireless rechargeable sensor networks, Springer, 2015.

[97] R. Lajara, J. Pérez-Solano, J. Pelegrí-Sebastiá, Predicting the batteries' state of health in wireless sensor networks applications, IEEE Transactions on Industrial Electronics 65 (11) (2018) 8936–8945.

[98] Y. Luo, L. Pu, G. Wang, Y. Zhao, Rf energy harvesting wireless communications: Rf environment, device hardware and practical issues, Sensors 19 (13) (2019) 3010.

[99] H. Lin, D. Bai, D. Gao, Y. Liu, Maximum data collection rate routing protocol based on topology control for rechargeable wireless sensor networks, Sensors 16 (8) (2016) 1201.

[100] C. Zhao, H. Zhang, F. Chen, S. Chen, C. Wu, T. Wang, Spatiotemporal charging scheduling in wireless rechargeable sensor networks, Computer Communications 152 (2020) 155–170.

[101] D. Arivudainambi, S. Balaji, Optimal placement of wireless chargers in rechargeable sensor networks, IEEE Sensors Journal 18 (10) (2018) 4212–4222.

[102] P. Zhong, Y. Zhang, S. Ma, J. Gao, Y. Chen, An adaptive mac protocol for wireless rechargeable sensor networks, in: International Conference on Wireless Algorithms, Systems, and Applications, Springer, 2017, pp. 244–252.

[103] Y. El Assari, Energy-efficient multi-hop routing with unequal clustering approach for wireless sensor networks, International Journal of Computer Networks & Communications (IJCNC) Vol 12.

[104] T. Biswas, S. Kumar, T. Singh, K. Gupta, D. Saxena, A comparative analysis of unequal clustering-based routing protocol in wsns, in: Soft Computing and Signal Processing, Springer, 2019, pp. 53–62.

[105] D. Agrawal, S. Pandey, Fuca: Fuzzy-based unequal clustering algorithm to prolong the lifetime of wireless sensor networks, International Journal of Communication Systems 31 (2) (2018) e3448.

[106] R. Ranganathan, B. Somanathan, K. Kannan, Fuzzy-based cluster head amendment (fcha) approach to prolong the lifetime of sensor networks, Wireless Personal Communications 110 (3) (2020) 1533–1549.

[107] S. Phoemphon, C. So-In, P. Aimtongkham, T. Nguyen, An energy-efficient fuzzy-based scheme for unequal multihop clustering in wireless sensor networks, JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING.

[108] P. S. Rao, H. Banka, Novel chemical reaction optimization based unequal clustering and routing algorithms for wireless sensor networks, Wireless Networks 23 (3) (2017) 759–778.

[109] N. Mazumdar, H. Om, Distributed fuzzy approach to unequal clustering and routing algorithm for wireless sensor networks, International journal of Communication systems 31 (12) (2018) e3709.

[110] S. Singh, P. Kumar, J. Singh, M. Alryalat, An energy efficient routing using multi-hop intra clustering technique in wsns, in: IEEE Region 10 Conference (TENCON), 2017, pp. 381–386.

[111] C. W. Wu, T. C. Chiang, L. C. Fu, An ant colony optimization algorithm for multi-objective clustering in mobile ad hoc networks, in: 2014 IEEE Congress on Evolutionary Computation (CEC), 2014, pp. 2963–2968.

[112] C. Caballero-Gil, P. Caballero-Gil, J. Molina-Gil, Self-organized clustering architecture for vehicular ad hoc networks, International Journal of Distributed Sensor Networks 11 (8) (2015) 384869.

[113] V. Mahima, A. Chitra, A novel energy harvesting: cluster head rotation scheme (eh-chrs) for green wireless sensor network (gwsn), Wireless Personal Communications 107 (2) (2019) 813–827.

[114] S. Ghosh, S. Mondal, U. Biswas, Enhanced pegasis using ant colony optimization for data gathering in wsn, in: 2016 International Conference on Information Communication and Embedded Systems (ICICES), IEEE, 2016, pp. 1–6.

[115] J. Albath, M. Thakur, S. Madria, Energy constraint clustering algorithms for wireless sensor networks, Ad Hoc Networks 11 (8) (2013) 2512–2525.

[116] M. Mezghani, An efficient multi-hops clustering and data routing for wsns based on khalimsky shortest paths, Journal of Ambient Intelligence and Humanized Computing 10 (4) (2019) 1275–1288.

[117] I. Turgut, Analysing multi-hop intra-cluster communication in cluster-based wireless sensor networks, Natural and Engineering Sciences.

[118] V. Deshpande, A. Patil, Energy efficient clustering in wireless sensor network using cluster of cluster heads, in: international conference on wireless and optical communications networks (WOCN), IEEE, 2013, pp. 1–5.

[119] L. Xu, M. O'Grady, G. O'Hare, R. Collier, Reliable multihop intra-cluster communication for wireless sensor networks, in: International Conference on Computing, Networking and Communications (ICNC), IEEE, 2014, pp. 858–863.

[120] A. Essa, A. Al-Dubai, I. Romdhani, M. Eshaftri, A new dynamic weight-based energy efficient algorithm for sensor networks, in: Smart grid inspired future technologies, Springer, 2017, pp. 195–203.

[121] S. Preeth, R. Dhanalakshmi, R. Kumar, P. Shakeel, An adaptive fuzzy rule based energy efficient clustering and immune-inspired routing protocol for wsn-assisted iot system, Journal of Ambient Intelligence and Humanized Computing (2018) 1–13.

[122] G. Han, L. Zhang, Wpo-eecrp: energy-efficient clustering routing protocol based on weighting and parameter optimization in wsn, Wireless Personal Communications 98 (1) (2018) 1171–1205.

[123] S. Ebrahimi Mood, M. Javidi, Energy-efficient clustering method for wireless sensor networks using modified gravitational search algorithm, Evolving systems 11 (4) (2020) 575–587.

[124] I. Akyildiz, X. Wang, Cross-layer design in wireless mesh networks, IEEE Transactions on Vehicular Technology 57 (2) (2008) 1061–1076.

[125] O. Landsiedel, E. Ghadimi, S. Duquennoy, M. Johansson, Low power, low delay: opportunistic routing meets duty cycling, in: ACM/IEEE 11th International Conference on Information Processing in Sensor Networks (IPSN), 2012, pp. 185–196.

[126] I. Rhee, A. Warrier, J. Min, L. Xu, Drand: Distributed randomized tdma scheduling for wireless ad hoc networks, IEEE Transactions on Mobile Computing 8 (10) (2009) 1384–1396.

[127] L. Louail, V. Felea, Latency optimization through routing-aware time scheduling protocols for wireless sensor networks, Computers & Electrical Engineering 56 (2016) 418–440.

[128] K. Heurtefeux, F. Maraninchi, F. Valois, Areacast: A cross-layer approach for a communication by area in wireless sensor networks, in: 17th IEEE International Conference on Networks, 2011, pp. 112–117.

[129] M. S. Hefeida, T. Canli, A. Khokhar, Cl-mac: A cross-layer mac protocol for heterogeneous wireless sensor networks, Ad Hoc Networks 11 (1) (2013) 213–225.

[130] C. Chou, K. Chuang, Colanet: a cross-layer design of energy-efficient wireless sensor networks, in: Systems Communications (ICW'05, ICHSN'05, ICMCS'05, SENET'05), IEEE, 2005, pp. 364–369.

[131] S. Schmid, R. Wattenhofer, Algorithmic models for sensor networks, in: Proceedings 20th IEEE International Parallel Distributed Processing Symposium (IPDPS), Rhodes Island, Greece, 2006, pp. 450–461.

[132] D. Zhai, R. Zhang, L. Cai, B. Li, Y. Jiang, Energy-efficient user scheduling and power allocation for noma-based wireless networks with massive iot devices, IEEE Internet of Things Journal 5 (3) (2018) 1857–1868.

[133] J. OMadadhain, D. Fisher, P. Smyth, S. White, Y. Boey, Analysis and visualization of network data using jung, Journal of Statistical Software 10 (2) (2005) 1–35.

[134] K. Cengiz, T. Dag, Energy aware multi-hop routing protocol for wsns, IEEE access 6 (2017) 2622–2633.

[135] A. S. Toor, A. Jain, Energy aware cluster based multi-hop energy efficient routing protocol using multiple mobile nodes (meacbm) in wireless sensor networks, AEU-International Journal of Electronics and Communications 102 (2019) 41–53.

[136] S. Sert, A. Alchihabi, A. Yazici, A two-tier distributed fuzzy logic based protocol for efficient data aggregation in multihop wireless sensor networks, IEEE Transactions on Fuzzy Systems 26 (6) (2018) 3615–3629.

[137] K. Kumaresan, S. Kalyani, Energy efficient cluster based multilevel hierarchical routing for multi-hop wireless sensor network, Journal of Ambient Intelligence and Humanized Computing 12 (3) (2021) 3821–3830.

[138] T. Ahmed, M. Abdur, A. Zaman, M. Amjad, A scalable k-hop clustering algorithm for pseudolinear manet, International Journal of Computer Applications 180 (2018) 62–68.

[139] M. Elhoseny, A. Farouk, N. Zhou, M. M. Wang, S. Abdalla, J. Batle, Dynamic multi-hop clustering in a wireless sensor network: Performance improvement, Wireless Personal Communications 95 (4) (2017) 3733–3753.

[140] R. A. Sadek, Hybrid energy aware clustered protocol for iot heterogeneous network, Future Computing and Informatics Journal 3 (2) (2018) 166–177.

[141] A. El Oualkadi, F. Choubani, A. El Moussati, Proceedings of the Mediterranean Conference on Information & Communication Technologies 2015: MedCT 2015, Vol. 2, 2016.

[142] L. Addario-Berry, N. Broutin, C. Goldschmidt, G. Miermont, The scaling limit of the minimum spanning tree of the complete graph, The Annals of Probability 45 (5) (2017) 3075–3144.

[143] A. A. Abbasi, M. Younis, A survey on clustering algorithms for wireless sensor networks, Computer communications 30 (14-15) (2007) 2826–2841.

[144] B. Neggazi, N. Guellati, M. Haddad, H. Kheddouci, Efficient self-stabilizing algorithm for independent strong dominating sets in arbitrary graphs, International Journal of Foundations of Computer Science 26 (06) (2015) 751–768.

[145] L. Xu, M. J. O'Grady, G. M. O'Hare, R. Collier, Reliable multihop intra-cluster communication for wireless sensor networks, in: 2014 International Conference on Computing, Networking and Communications (ICNC), 2014, pp. 858–863.

[146] J. Xu, W. Liu, F. Lang, Y. Zhang, C. Wang, Distance measurement model based on rssi in wsn, Wireless Sensor Network 2 (8) (2010) 606.

[147] D. Kim, M. Jung, Data transmission and network architecture in long range low power sensor networks for iot, Wireless Personal Communications 93 (1) (2017) 119–129.

[148] X. Rao, P. Yang, Y. Yan, H. Zhou, X. Wu, Optimal recharging with practical considerations in wireless rechargeable sensor network, IEEE Access 5 (2017) 4401–4409.

[149] E. Siow, T. Tiropanis, W. Hall, Analytics for the internet of things: A survey, ACM Computing Surveys (CSUR) 51 (4) (2018) 1–36.

[150] M. Shuai, J. Modi, T. Peng, S. Chengyi, W. Jianbo, . Jun, D. Tao, S. Wen, Temperature effect and thermal impact in lithium-ion batteries: A review, Progress in Natural Science: Materials International 28 (6) (2018) 653–666.

[151] R. Fitria, Y. Leny, S. Imam, R. Fatmawati, Carbon rod of zinc-carbon primary battery waste as a substrate for cds and tio2 photocatalyst layer for visible light driven photocatalytic hydrogen production, Journal of Environmental Chemical Engineering 5 (3) (2017) 2251–2258.

[152] G. E. Blomgren, The development and future of lithium ion batteries, Journal of The Electrochemical Society 164 (1) (2016) A5019.

[153] B. Liao, Y. Lei, L. Chen, G. Lu, H. Pan, Q. Wang, Effect of the la/mg ratio on the structure and electrochemical properties of laxmg3xni9 (x=1.6–2.2) hydrogen storage electrode alloys for nickel–metal hydride batteries, Journal of Power Sources 129 (2) (2004) 358–367.

[154] K. Abraham, How comparable are sodium-ion batteries to lithium-ion counterparts?, ACS Energy Letters 5 (11) (2020) 3544–3547.

[155] K. Ergun, R. Ayoub, P. Mercati, T. Rosing, Dynamic optimization of battery health in iot networks, in: IEEE 37th International Conference on Computer Design (ICCD), 2019, pp. 648–655.

[156] B. Xu, A. Oudalov, A. Ulbig, G. Andersson, D. Kirschen, Modeling of lithium-ion battery degradation for cell life assessment, IEEE Transactions on Smart Grid 9 (2) (2018) 1131–1140.

[157] K. Uddin, S. Perera, W. Widanage, L. Somerville, J. Marco, Characterising lithium-ion battery degradation through the identification and tracking of electrochemical battery model parameters, Batteries 2 (2) (2016) 13.

[158] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, P. Van den Bossche, Critical review of state of health estimation methods of li-ion batteries for real applications, Renewable and Sustainable Energy Reviews 56 (2016) 572–587.

[159] P. Ding, J. Holliday, A. Celik, Distributed energy-efficient hierarchical clustering for wireless sensor networks, in: International conference on distributed computing in sensor systems, 2005, pp. 322–339.

[160] S. Ma, M. Jiang, P. Tao, C. Song, J. Wu, J. Wang, T. Deng, W. Shang, Temperature effect and thermal impact in lithium-ion batteries: A review, Progress in Natural Science: Materials International 28 (6) (2018) 653–666.

[161] L. Raijmakers, D. Danilov, R.-A. Eichel, P. Notten, A review on various temperature-indication methods for li-ion batteries, Applied energy 240 (2019) 918–945.

[162] C. Forgez, D. V. Do, G. Friedrich, M. Morcrette, C. Delacourt, Thermal modeling of a cylindrical lifepo4/graphite lithium-ion battery, Journal of Power Sources 195 (9) (2010) 2961–2968.

[163] P. Taheri, M. Bahrami, Temperature rise in prismatic polymer lithium-ion batteries: An analytic approach, SAE International Journal of Passenger Cars - Electronic and Electrical Systems 5 (2012) 164176.

[164] M. Bhuiyan, G. Wang, A. Vasilakos, Local area prediction-based mobile target tracking in wireless sensor networks, IEEE Transactions on Computers 64 (7) (2014) 1968–1982.

ملخص:

تركز أطروحة الدكتوراه هذه على تصميم ونمذجة طرق تخصيص وتجميع القنوات الجديدة للشبكات اللاسلكية واسعة النطاق. بالإضافة إلى ذلك ، يركز هذا العمل على تحسين عمر البطاريات القابلة لإعادة الشحن في بروتوكولات الشبكة اللاسلكية لتقليل استبدال المواد وزيادة متانة الشبكة على المدى الطويل. في هذا السياق، نبدأ بتحليل الأدبيات لتحديد بعض المشاكل المفتوحة أو الحلول المقترحة التي لا يزال يتعين تحسينها. نركز أولاً على طبقة MAC وبروتوكولات التوجيه ونقترح نهجًا جديدًا لجدولة TDMA بين الطبقات. يأخذ هذا الأخير المعلومات من شجرة التوجيه لإدارة الوصول إلى القناة اللاسلكية وتقليل زمن انتقال الاتصال. بعد ذلك، ركزنا على إدارة طوبولوجيا الشبكة واقترحنا طرقًا مختلفة لتجميع الشبكات المتكافئ وغير المتكافئ. بعد ذلك، ركزنا على تحسين الطاقة على المدى الطويل للأجهزة اللاسلكية القابلة لإعادة الشحن واقترحنا طرقًا أصلية جديدة في هذا السياق الأصلي. الهدف الرئيسي من الأساليب المقترحة هو تقليل زمن انتقال الاتصال وتقليل الحِمل على الشبكة وزيادة عمر الشبكة على المدى القصير والطويل.

مفاتيح: إنترنت الأشياء، شبكة الاستشعار اللاسلكية، هيكلة الشبكات، قنوات الاتصال، تحسين استهلاك الطاقة على المدى الطويل.

Abstract:

The main goal of this thesis focuses on the design and modeling of new channel allocation and clustering approaches for large scale wireless networks. Additionally, this works focus on the optimization of rechargeable battery lifespan in wireless networking protocols to reduce the material substitution and increase the network durability in long-term. In this context, we start with an analysis of the literature to identify certain open problems or whose proposed solutions that still to improve. We firstly concentrate on the MAC and routing layer protocols and we propose a new cross layer TDMA scheduling approaches. This latter considers the information of the routing tree to manage the wireless channel access and reduce the communication latency. Then, we focus on the network topology management and we proposed different approaches for the equal and unequal network clustering. Afterwards, we focus on the long-term energy optimization of rechargeable wireless devices and we proposed new original approaches in this original context. The main objective of the proposed approaches is to decrease the communication latency, reduce the network overhead, and increase the network lifespan in the short and long-term.

Keywords: IoT, WSN, Dynamic Clustering, Network Connectivity, Rechargeable Battery Lifetime, Energy-aware Protocols, Long-term Energy Optimization.

Résumé:

Cette thèse de doctorat porte sur la conception et la modélisation de nouvelles approches d'allocation de canaux et de clustering pour les réseaux sans fil à grande échelle. De plus, ces travaux se concentrent sur l'optimisation de la durée de vie des batteries rechargeables dans les protocoles de réseau sans fil afin de réduire la substitution de matériaux et d'augmenter la durabilité du réseau à long terme. Dans ce cadre, nous commençons par une analyse de la littérature pour identifier certains problèmes ouverts ou dont les solutions proposées restent à améliorer. Nous nous concentrons tout d'abord sur les protocoles de couche MAC et de routage et nous proposons une nouvelle approche d'ordonnancement TDMA inter couche. Ce dernier considère les informations de l'arbre de routage pour gérer l'accès au canal sans fil et réduire la latence de communication. Ensuite, nous nous sommes concentrés sur la gestion de la topologie du réseau et nous avons proposé différentes approches pour le clustering réseau homogène et hétérogène. Par la suite, nous nous sommes concentrés sur l'optimisation énergétique à long terme des dispositifs sans fil rechargeables et nous avons proposé de nouvelles approches originales dans ce contexte original. L'objectif principal des approches proposées est de diminuer la latence de communication, de réduire la surcharge du réseau et d'augmenter la durée de vie du réseau à court et à long terme.

Mots clés: Internet des objets, Réseau de capteurs sans fil, Structuration des réseaux, Canaux de communication, Efficacité énergétique, Optimisation énergétique à long terme.